

UNIVERSITÉ DE SHERBROOKE-ECOLE DE GESTION

DÉTERMINATION DES PRIORITÉS À CONSIDÉRER LORS DE L'ÉLABORATION D'UN CODE D'ÉTHIQUE ET DE DÉONTOLOGIE DANS LE DOMAINE DE L'IA&A

Mémoire dans le cadre de la maîtrise en stratégie de l'intelligence d'affaires
cheminement recherche

Réalisé par Chimène Flore FADJIE NGAYAP
sous la direction du Professeur Daniel Chamberland-Tremblay
04/12/2019

Baliser la pratique d'un métier à travers un code d'éthique et de déontologie contribue entre autres à augmenter la confiance entre les partenaires d'affaires ainsi que toutes les parties prenantes, à légitimer le travail des professionnels et enfin à garantir l'image de marque de l'entreprise. Le présent document présente le résultat d'une étude qui permet de définir les bases conceptuelles d'un code d'éthique et de déontologie ainsi que son mode de gestion dans le cadre de la profession de l'intelligence d'affaires et analytique et précisément pour ce qui concerne l'exploitation des données.

SOMMAIRE :

Cette recherche vise à poser les bases conceptuelles pour l'élaboration d'un code d'éthique adapté, répondant aux impératifs des professionnels du domaine de l'intelligence d'affaires et analytique en l'absence de balises actuelles encadrant l'exploitation éthique des données malgré l'engouement de plus en plus croissant pour les données. L'étude examine deux éléments. Premièrement, les principes fondateurs d'un code d'éthique dédié à l'intelligence d'affaires et analytique (IA&A) selon la perspective des professionnels. Deuxièmement, le mode de gestion à privilégier par ces mêmes professionnels. Les professionnels interrogés dans cette étude sont divisés en deux catégories selon le rôle qu'ils jouent vis-à-vis de la donnée. Il s'agit des scientifiques de données et des gestionnaires d'équipes d'IA&A.

Étant donné le volume croissant d'informations recueillies et analysées au quotidien avec ou sans le consentement des individus, le comportement des professionnels doit respecter dans ce contexte le code d'éthique et de déontologie de l'entreprise afin de garantir pour les organisations leur image de marque, garder une relation privilégiée avec leur clientèle et aider à garder leur revenu, pour ne nommer que cela.

Cette étude est donc une première étape vers la définition et la mise en place d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données. DELPHI a été utilisée comme méthode de recherche principale dans cette étude. 16 professionnels ont été sollicités au premier tour. Un questionnaire contenant 5 questions leur a été acheminé. 22 formulaires ont été complétés et retournés par les professionnels à l'issue de la première itération. Un deuxième questionnaire comportant des éléments de réponses des 22 professionnels ayant retourné leur formulaire du premier tour a été envoyé ; ceci dans le but de chercher une convergence d'opinions de la part des professionnels. Finalement, 14 formulaires ont été retournés ; ce qui correspond à taux de réponses satisfaisant selon les critères de la méthodologie de recherche adoptée. Le nombre de participants minimum selon la méthode DELPHI se situant entre 7 et 18 experts.

Les résultats démontrent que :

- Les éléments anonymat, confidentialité, secret, intégrité, transparence et sécurité sont bien compris et sont considérés importants à l'exploitation éthique des données selon la majorité des répondants.
- Le mode de gestion privilégié varie selon le rôle des professionnels vis-à-vis des données, les gestionnaires sont en grande majorité en faveur d'une gestion centralisée tandis que les avis sont partagés au sein des scientifiques de données.
- Les professionnels sont unanimement en faveur d'un mécanisme de contrôle interne et automatisé et ouvert malgré des réticences à un organisme étatique pour un contrôle externe.

Les implications et les contributions de la recherche sont soulignées dans le présent document.

TABLE DE MATIÈRES

SOMMAIRE :	2
REMERCIEMENT	6
CHAPITRE 1 : INTRODUCTION	7
CHAPITRE 2 : REVUE DE LITTÉRATURE	9
1. L'intelligence d'affaires et l'analytique	9
a) L'intelligence d'affaires	10
b) l'analytique ou science de données	13
2. Le scientifique de données	14
3. Modèle de maturité en intelligence d'affaires	16
4. L'éthique, la morale et la déontologie	20
a) L'éthique	21
b) L'éthique versus la morale	22
c) L'éthique versus la légalité	23
d) L'éthique versus la déontologie	24
5. L'éthique en intelligence d'affaires	26
6. Quelques enjeux éthiques en intelligence d'affaires	30
7. Ancrages déontologiques	34
a) Codes d'éthique et de déontologie	34
❖ Cas des ressources humaines	34
❖ Cas de la profession comptable	35
❖ Cas du marketing	36
❖ Cas de l'intelligence d'affaires	36
Conclusion	38
CHAPITRE 3 : PROBLÉMATIQUE	39
1. Mise en contexte de la recherche	39
2. Importance du sujet de recherche	43

3. Problématique de recherche	45
4. Objectifs de recherche et résultats anticipés.....	47
CHAPITRE 4 : MÉTHODOLOGIE.....	49
1. Choix de la méthode	49
2. Description de la méthode DELPHI	50
3. Avantages de la méthode.....	54
4. Limites de la méthode.....	55
5. Conception de l'enquête pour l'étude	56
a) la Procédure de sélection des experts	57
b) Processus d'administration du questionnaire.....	59
CHAPITRE 5 : RÉSULTATS ET DISCUSSION	63
1. Description des résultats.....	63
a) Résultats des professionnels en exploitation des données	63
b) Résultat des gestionnaires d'équipes de scientifiques de données.....	76
2. Discussion.....	88
a) Comparaison des résultats proposés par les deux groupes des répondants au premier tour	88
b) Conclusions importantes.....	101
3. Contributions.....	102
a) Contributions pratiques	103
b) Contributions théoriques	104
4. Limites de la recherche	105
5. Recherches futures.....	106
CHAPITRE 6 : CONCLUSION	108
BIBLIOGRAPHIE:	110
LEXIQUE :	117
ANNEXES	118

Table des illustrations

Figure 1	12
Figure 2	14
Figure 3	17
Figure 4	20
Figure 5	40
Tableau 1	53
Tableau 2	53
Tableau 3	65
Tableau 4	89

REMERCIEMENT

Merci, Adonaï, tu es digne d'honneur et tu rends toute chose possible.

Merci Professeur John Ingham, vous avez su me transmettre par votre passion et votre professionnalisme, le désir de faire de la recherche. Changer de plan d'étude n'a pas été difficile pour moi malgré la charge de travail que cela représentait.

Merci Professeur Daniel Chamberland-Tremblay. Professeur, vous êtes pour moi un modèle; votre amour du travail bien fait, votre aisance à transmettre votre savoir et votre savoir-faire, votre capacité à encourager et à amener l'étudiant à se dépasser. Merci professeur pour votre soutien, vos conseils, votre temps, votre implication dans ce projet, vos instructions et recommandations.

Merci, M. Pierre Cusson, vous m'avez motivé à entreprendre une formation en intelligence d'affaires et analytique et aujourd'hui, cela aboutit à une maîtrise.

Merci aux experts qui ont participé à l'enquête en qualité de répondants. Votre participation a contribué à réaliser ce travail, mais aussi à améliorer le champ de recherche.

Merci aux membres de jury pour le travail de lecture et d'évaluation du présent mémoire.

Merci à ma famille et à mes amis pour votre soutien indéfectible, vos encouragements, vos prières et votre présence.

CHAPITRE 1 : INTRODUCTION

Les dernières décennies ont imposé des changements importants dans l'environnement d'affaires des organisations petites et grandes. Par exemple, la mondialisation des affaires engendre une concurrence accrue et une pression vers la productivité et l'efficience. L'avènement et la croissance d'Internet et du Web ont aussi transformé les relations avec les clients, les fournisseurs, les investisseurs et les autres parties prenantes. Un des impacts les plus visibles de ce changement pour l'organisation est la quantité et la diversité des données auxquelles elle a accès. Dans ce contexte, il n'est pas surprenant de répertorier aujourd'hui un grand nombre de technologies et de pratiques qui visent à exploiter et valoriser ces données.

L'intelligence d'affaires tout comme l'analytique, vise à outiller des professionnels pour mieux exploiter et valoriser les données.

En effet, diverses recherches ont mis en évidence l'existence d'un lien étroit entre la performance d'une entreprise en termes de rentabilité, de revenus et de rendement pour les actionnaires, et son utilisation de l'analytique (Evans, 2012). Les entreprises performantes (celles qui surpassent leurs concurrents) sont trois fois plus susceptibles d'être plus sophistiquées dans leur utilisation de l'analytique (techniques et procédés d'exploitations de données) que les entreprises moins performantes. Elles sont également susceptibles d'affirmer que leur utilisation de l'analytique les différencie de leurs concurrents (Davenport et Harris, 2007 ; Hopkins et al., 2010 ; Evans, 2012).

L'augmentation des capacités d'exploitation des données numériques offre aux gestionnaires de nouveaux outils pour percevoir, comprendre, décider et agir. Si dans bien des cas, l'utilisation des données permet à l'organisation d'améliorer l'expérience du client, du fournisseur, de l'employé ou d'une autre partie prenante, le rythme effréné du développement des capacités de collecte et d'analyse de données ne permet pas à l'ensemble de la population de bien comprendre les enjeux sous-jacents ou aux législateurs de baliser les pratiques du domaine. Ce décalage engendre des dérives importantes notamment le non-respect de la vie privée des individus, le bris de confidentialité, l'absence de consentement éclairé dans certains usages de technologies, l'image de marque ternie, pour ne citer que ceux-là. Qui plus est, les capacités accrues de stockage et d'intégration des données de sources diverses permettent aux

organisations d'établir des bases de données hautement stratégiques, mais posant des risques pour les individus dont le profil y est stocké (fuite de données, vol d'identité).

Tout n'est pas idéal, les risques perçus sont tout à fait réels. L'intelligence d'affaires s'intéresse principalement à quelles données acquérir, comment les acquérir et comment les transformer en informations utiles à une prise de décision éclairée ?

Il y a des trous, et contrairement aux autres disciplines des affaires telles que la comptabilité, le marketing, ou encore les ressources humaines, on constate que **la discipline de l'intelligence d'affaires manque un cadre déontologique qui permet à ses professionnels d'agir au quotidien en respectant les balises éthiques**. Ce constat sonne une cloche d'alarme. Dans le milieu académique où des chercheurs comme Thomann et Wells (2013) ont réfléchi sur la gestion de l'éthique en intelligence d'affaires et attiré l'attention sur les thématiques qui soulèvent des enjeux éthiques. Au niveau pratique, on se demande comment baliser les pratiques des spécialistes. Dans la dernière année, on se pose la question pour l'intelligence artificielle. Mais ces questions restent toujours en suspens pour l'analyste humain.

La conduite de cette étude empirique et exploratoire aborde la thématique de la **détermination des priorités à considérer lors de l'élaboration d'un code d'éthique et de déontologie dans le domaine de l'intelligence d'affaires et de l'analytique**.

Ce mémoire compte cinq sections. La première section présente une revue de littérature qui décrit les concepts liés à l'éthique, à la science des données et à l'intelligence d'affaires. La deuxième section établit la question de recherche. Ici, on fait la mise en contexte de la pertinence et nouveauté de l'objet de cette recherche ; on énonce la problématique et les objectifs de recherche. La troisième section détaille la méthodologie notamment le choix de la méthode, sa description ainsi que ses limites. La quatrième section traite les résultats de recherche où on décrit la collecte des données ainsi que les résultats, puis on met en perspective les résultats et présente les conclusions importantes. Finalement, la conclusion permet de revenir sur l'intégrabilité de la recherche, les découvertes faites ainsi que les limites du travail tout en présentant les perspectives à venir.

CHAPITRE 2 : REVUE DE LITTÉRATURE

Le chapitre précédent introduit le présent travail de recherche et fait état de la perspective professionnelle de l'éthique déontologique en intelligence d'affaires et analytique (IA&A). Les pratiques douteuses de collecte de données personnelles, les fuites de données en entreprise et l'émergence de l'intelligence artificielle en affaires ont contribué à augmenter la conscience collective. On constate cependant que la formation en éthique chez les professionnels des données est limitée, voire inexistante. Ce portrait à lui seul démontre la pertinence et la nécessité de cette étude exploratoire auprès des professionnels du domaine.

Ce chapitre pose les cadres conceptuels et théoriques nécessaires à la bonne compréhension de cette étude. Dans un premier temps, l'intelligence d'affaires et l'analytique sont formellement définis. Dans un deuxième temps, les concepts d'éthique et de déontologie sont abordés et distingués. Enfin, la section se termine sur la notion de gouvernance éthique des données en entreprise.

1. L'intelligence d'affaires et l'analytique

L'intelligence d'affaires et l'analytique (IA&A) peut être définie à haut niveau comme le domaine de la gestion et de la valorisation des données de l'organisation pour appuyer la prise de décisions (Chen, Chiang et Storey, 2012 ; Foley et Guillemette, 2010). Elle trouve ses racines dans la gestion de bases de données (Chaudhuri et al., 2011 ; Turban et al., 2008 ; Watson et Wixom, 2007) et les systèmes d'aide à la décision (Foley et Guillemette, 2010). Elle s'inscrit à la frontière des systèmes d'information, de l'informatique et de la gestion.

Les termes intelligence d'affaires, science des données et analytique sont souvent utilisés comme un seul et même concept. Cependant, chacun des termes présente des spécificités qui lui sont propres. Les sections suivantes décrivent les concepts clés du domaine de façon à bien baliser la recherche.

a) L'intelligence d'affaires

Le terme « intelligence d'affaires » a été utilisé pour la première fois en 1989 par M. Howard Dressner, chercheur au Gartner Group, comme terme générique pour décrire des concepts et des méthodes visant à améliorer la prise de décision d'affaires en utilisant un support basé sur des faits (Negash et Gray, 2008). Le terme a trouvé un écho auprès des professionnels du domaine des systèmes d'aide à la décision, des fournisseurs technologiques et des gestionnaires en entreprise. Il a été largement adopté et a remplacé des termes tels que systèmes d'information de direction ou systèmes experts.

Depuis, la présence de l'intelligence d'affaires croît dans les organisations (Foley et Guillemette, 2010 ; Dussureaul, 2015). Les technologies de l'information plus puissantes et le besoin de prendre des décisions plus complexes dans un contexte d'affaires plus compétitif et plus évolutif favorisent l'adhésion à l'intelligence d'affaires.

Néanmoins, l'intelligence d'affaires est un terme évolutif qui désigne plusieurs réalités dans l'entreprise au fil des innovations technologiques et des modes de l'industrie (Negash et Gray, 2008). Power (2002) définit les systèmes d'intelligence d'affaires comme étant des systèmes d'aide à la décision (DSS) basés sur des données qui mettent l'accent sur l'analyse de grands volumes de données structurées et semi-structurées. Negash et Gray (2008) définissent l'intelligence d'affaires comme des systèmes qui combinent la collecte de données, le stockage des données et la gestion des connaissances avec l'analyse pour évaluer l'information complexe de l'entreprise et de la concurrence afin de la présenter aux planificateurs et aux décideurs, dans le but d'améliorer la rapidité et la qualité de l'apport au processus décisionnel. Foley et Guillemette (2010) élargissent la définition de l'intelligence d'affaires pour tenir compte des aspects de gestion du domaine. Ils intègrent la notion de culture et de politique. Ces derniers proposent cette définition formelle de l'intelligence d'affaires :

Business intelligence (BI) is a combination of processes, policies, culture, and technologies for gathering, manipulating, storing, and

analyzing data collected from internal and external sources, in order to communicate information, create knowledge, and inform decision making. BI helps report business performance, uncover new business opportunities, and make better business decisions regarding competitors, suppliers, customers, financial issues, strategic issues, products and services.

Les éléments « données », « analyse », « information » et « connaissances » sont communs à toutes les définitions offertes par les auteurs cités précédemment. En intelligence d'affaires, on distingue donc les données qui encodent les faits, l'information qui est issue de la mise en contexte des données et les connaissances qui découlent de l'utilisation de l'information dans l'analyse d'une situation. Cette hiérarchie correspond à celle proposée par Ackoff (1989).

L'intelligence d'affaires inclut également les architectures, les outils, les bases de données, les applications et les méthodologies sous-jacentes (Ee-Peng, Hsinchun, et Guoqing, 2012). En tant qu'approche centrée sur les données, l'intelligence d'affaires s'appuie fortement sur diverses technologies avancées de collecte, d'extraction et d'analyse de données (Turban et al., 2008 ; Watson et Wixom, 2007 ; Ee-Peng, Hsinchun, et Guoqing, 2012). Ces technologies sont collectivement connues sous le nom de « Business Analytics » (Ee-Peng, Hsinchun, et Guoqing, 2012). L'entrepotage de données étant souvent considéré comme le fondement de l'intelligence d'affaires. La conception de comptoirs de données et d'outils d'extraction, de transformation et de chargement (ETL) est essentielle à la conversion et à l'intégration des données propres à l'entreprise. L'interrogation de bases de données, le traitement analytique en ligne (OLAP) et les outils de visualisation avancés sont souvent adoptés par la suite pour explorer les caractéristiques importantes des données (Chaudhuri et al., 2011).

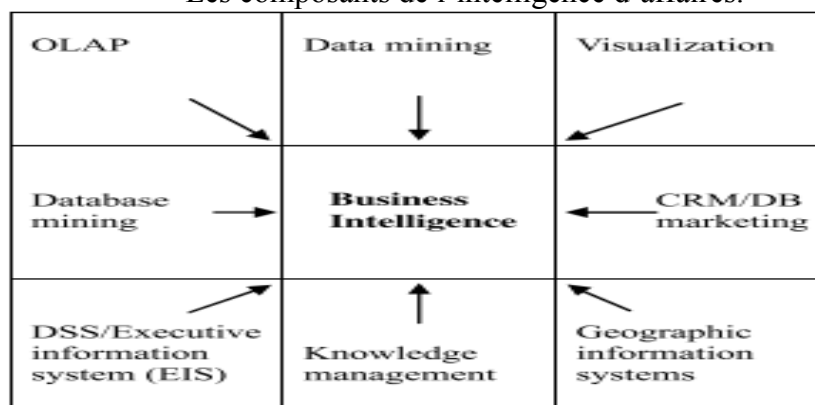
Les principaux objectifs de l'IA&A sont donc de permettre des échanges interactifs et faciles, favoriser l'accès à diverses données, permettre la manipulation et la transformation de ces données, et fournir aux gestionnaires et aux analystes d'affaires la capacité d'effectuer des analyses appropriées et prendre des décisions éclairées. (Turban et al., 2008 ; Wixom et al., 2011 ; Ee-Peng, Hsinchun, et Guoqing, 2012).

En guise de résumé, on peut retenir que l'intelligence d'affaires fait recours aux termes « données », « analyse », « information » et « connaissances ». Aussi elle comporte deux volets. D'abord, le volet axé sur les technologies de l'intelligence d'affaires, où il est question du maniement des données et des informations via les outils et procédés tels que les entrepôts de données, les processus ETL (*Extract, Transform, Load*), l'analyse multidimensionnelle, les cubes OLAP (*OnLine Analytical Processing*), le forage de données ou encore les tableaux de bord et cartes d'indicateurs (Chaudhuri et al., 2011 ; Dussureaul, 2015). Ensuite le second volet, celui de la création de savoir, la communication de l'intelligence et la prise de décision, cette partie fait référence au côté managérial de l'intelligence d'affaires et permet aux compagnies d'utiliser les informations et le savoir généré pour supporter leur processus de prise de décision opérationnelle et stratégique (Hannula et Pirttimaki, 2003 ; Dussureaul, 2015).

Enfin, l'intelligence d'affaires est le résultat d'une série d'innovations au fil des ans. Ses antécédents sont les systèmes d'aide à la décision décrits, par exemple, dans Alter (1980), les langages de planification financière de quatrième génération tels que IFPS (Gray, 1983), les systèmes d'information exécutive (Watson et al., 1992), l'entrepôt de données (Inmon, 1992), le traitement analytique en ligne (Codd, 1993), et le forage de données (Fayyad et al., 1996). Chaque itération étant plus sophistiquée que la précédente (Negash et Paul Gray, 2008). Aujourd'hui, les logiciels regroupent tous ces éléments en un seul système (Figure 1 : Negash et Gray, 2008).

Figure 1

Les composants de l'intelligence d'affaires.



Source: Negash et Gray (2008)

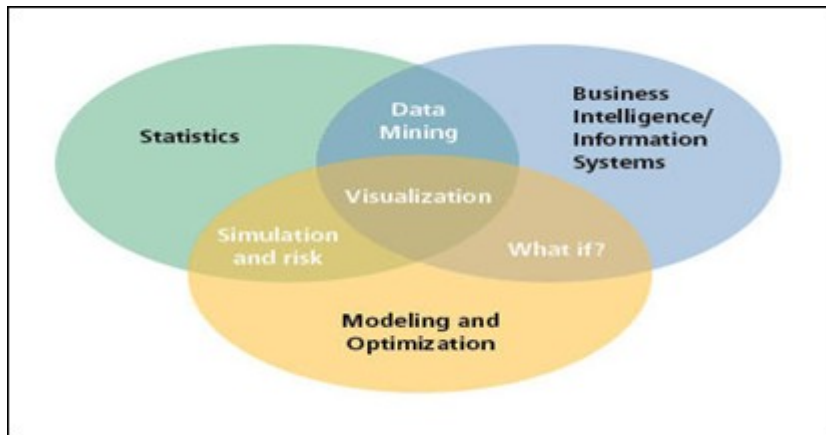
b) L'analytique ou science de données

L'intelligence d'affaires fournit aux entreprises de l'information pour les aider à accroître leur compétitivité (Negash et Gray, 2008). Mais l'information ne suffit pas à elle seule pour remplir cette mission. L'information doit être interprétée en fonction des objectifs stratégiques et tactiques de l'entreprise. En cela, Davenport et Harris (2007) soulignent le rôle de l'analytique comme celui de piloter les décisions et les actions managériales. Negash et Gray (2008) quant à eux considèrent l'analytique comme l'intrant de l'analyse pour les humains et les prises de décision automatisée. Ces auteurs laissent entrevoir l'apport de l'analytique comme guide dans la prise de décision, une fois que l'on a l'information. Liberatore et Luo (2010), définissent l'analytique comme « un processus de transformation des données en actions par l'analyse et la compréhension dans le contexte de la prise de décision organisationnelle et de la résolution de problèmes ». Russom (2011) va au-delà du processus et présente l'analytique comme un ensemble de techniques et de types d'outils connexes, incluant habituellement l'analyse prédictive, l'exploration de données, l'analyse statistique et le SQL complexe. La liste peut couvrir la visualisation des données, l'intelligence artificielle, le traitement du langage naturel et les capacités des bases de données qui supportent les analyses (telles que le MapReduce, les analyses dans les bases de données, les bases de données en mémoire, les banques de données en colonnes). Et enfin, Evans (2012) définit l'analytique comme « l'utilisation de données, de technologies de l'information, d'analyses statistiques, de méthodes quantitatives et de modèles mathématiques ou informatiques pour aider les gestionnaires à mieux comprendre leurs activités commerciales et à prendre de meilleures décisions fondées sur des faits ». Tout comme dans le cas l'intelligence d'affaires, on voit ici une évolution au fil du temps dans la définition du concept « analytique ». On remarque aussi que la définition est de plus en plus complète.

Dans la dernière définition, Evans (2012) présente l'analytique comme une convergence de trois disciplines qui ont été enseignées et utilisées depuis longtemps : les statistiques, les systèmes d'information et de veille économique, et la modélisation et l'optimisation. Pour cet auteur, la figure 2 montre les relations et les synergies qui définissent l'analytique.

Figure 2

Une perspective de l'analytique



Source: Evans (2012)

D'après la figure 2, l'intelligence d'affaires (en anglais *Business Intelligence*) fait partie intégrante de l'analytique. Par contre, on constate dans la figure 1 que l'analytique est un composant de l'intelligence d'affaires. Qu'en est-il exactement ? Plusieurs auteurs s'accordent pour dire que l'intelligence a évolué de son origine traditionnelle pour dépendre de l'analytique. Aussi, dans les paragraphes précédents, on a énoncé que l'analytique a besoin des données fournies par l'intelligence d'affaires. Il ressort donc de cela que les deux concepts se complètent et dépendent fortement l'un de l'autre. C'est sans doute pour cette raison que E-Peng, Hsinchun, et Guoqing (2012), suggèrent d'adopter **l'intelligence d'affaires et analytique** comme un **terme combiné** de préférence. Dans les lignes suivantes, on adoptera la proposition de ces auteurs et de ce fait, on fera référence à l'intelligence d'affaires et l'analytique et non plus à l'un ou l'autre concept de manière séparée.

2. Le scientifique de données

Le paragraphe précédent définit la science de données et l'intelligence d'affaires. Leurs origines et leurs pratiques courantes. Le paragraphe aboutit à la conclusion que ces deux concepts se complètent et dépendent fortement l'une de l'autre de sorte qu'on doit les utiliser comme un seul et même concept combiné sous l'appellation de l'intelligence d'affaires et analytique (IA&A). Il est question dans cette section de

définir qui est « le scientifique de données » ; surtout que la population qui composera les répondants à l'enquête menée par cette étude est faite de professionnels intégrant la caractéristique de scientifique de données.

En 1991, le statisticien Moore emploie le mot « science de données » pour la première fois, pour définir la statistique (Moore, 1991). Cette définition est reprise plus tard par d'autres auteurs, dont James T. McClave et al. en 1997. En 2013, Vasant Dhar dans son article intitulé « Data Science and Prediction » élargit le terme et sans faire référence seulement à la statistique, il définit la science des données comme étant l'étude de l'extraction des connaissances à partir d'un ensemble de données à travers des compétences intégrées englobant les mathématiques : la statistique principalement, la théorie de l'information et la technologie de l'information : dont le traitement de signal, des modèles probabilistes, l'apprentissage automatique, l'apprentissage statistique, la programmation informatique, l'ingénierie de données, la reconnaissance de formes et l'apprentissage, la visualisation, l'analytique prophétique, la modélisation d'incertitude, le stockage de données, la compression de données et le calcul à haute performance (Dhar, 2013).

Partant de cette dernière définition, c'est à juste titre qu'on pourrait parler d'un scientifique de données comme une personne qui pratique la science des données et qui est responsable du croisement des données de l'entreprise avec celles mises à disposition via les services Web et autres canaux digitaux (téléphone mobile, montre intelligente et autres objets connectés ...) dans le but de donner du sens à ces données et en extraire de la valeur afin d'aider l'entreprise à prendre des décisions stratégiques ou opérationnelles (Lesparre, 2019). Le scientifique des données est donc appelé à concevoir des modèles et algorithmes pour collecter, stocker, traiter et restituer les données et c'est cette forme de « créativité » qui le distingue du pur statisticien (Lesparre, 2019). En effet, il est capable d'imaginer de nouveaux modèles d'analyse pour traiter des données brutes et hétérogènes qui ne peuvent pas être analysées à l'aide d'outils classiques de gestion de bases de données.

Traditionnellement, le scientifique de données doit combiner une triple compétence :

- Expertise statistique et informatique,
- Connaissance des bases de données et de l'informatique,
- Expérience métier dans leur secteur d'activité (marketing, finance, aéronautique par exemple).

Cependant, lors de la recherche et l'identification des répondants, on a remarqué que beaucoup de personnes s'auto qualifient de scientifique de données dès la fin de leurs études universitaires; on comprend qu'à ce stade, on est loin de la définition traditionnelle du scientifique de données dans la mesure où l'étudiant qui vient d'obtenir son diplôme n'a pas encore eu le temps de développer la dernière compétence ci-dessus énumérée à savoir « expérience métier dans le secteur d'activité. »

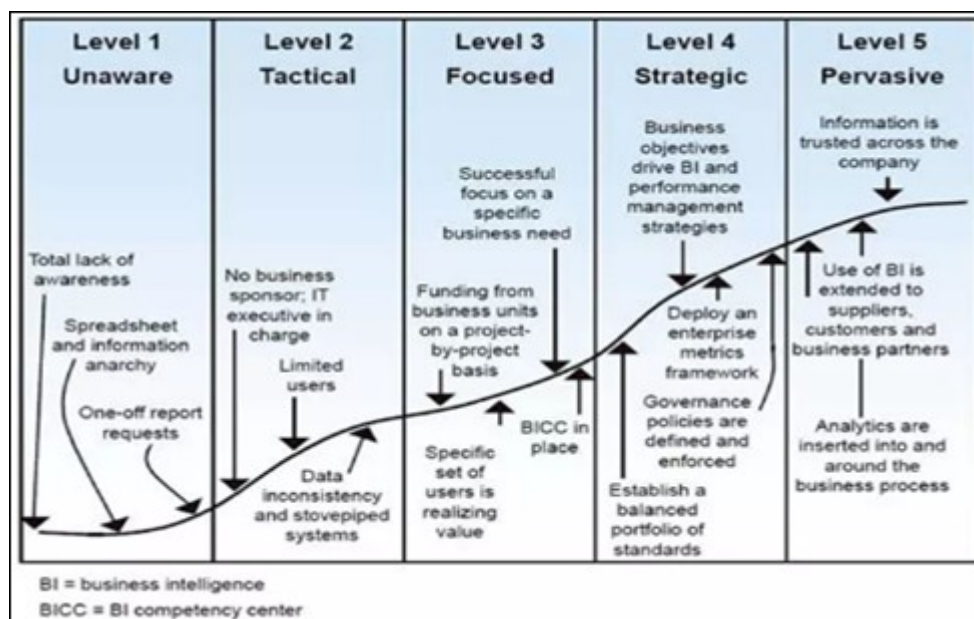
Est-ce que cette observation prépare à une future génération de scientifique de données où la connaissance métier en rapport avec le secteur d'activité ne sera plus nécessaire ? Cette problématique reste ouverte puisqu'elle n'est pas l'objet direct de la présente étude.

Cependant, on pense qu'il est important d'avoir une bonne connaissance du métier pour mieux interpréter les résultats des différentes analyses, raison pour laquelle des critères ont été définis et seront présentés plus loin lors du choix des répondants.

3. Modèle de maturité en intelligence d'affaires

Après avoir présenté l'intelligence d'affaires et l'analytique d'une part, et défini le scientifique de données d'autre part, il convient dans cette section de présenter le modèle de maturité en intelligence d'affaires et l'analytique. On en parle ici, car l'évolution de l'intelligence d'affaires traditionnelle vers l'IA&A est aussi une conséquence de la maturité de la discipline de base qu'est l'intelligence d'affaires comme vont le montrer les lignes suivantes. L'autre raison est que les modèles de maturité selon Oueslati (2018) sont utilisés pour guider le processus de transformation d'une organisation à partir d'une situation initiale vers un stade ciblé. On dénote selon cet auteur, l'existence de plusieurs modèles de maturité qui décrivent, expliquent et évaluent le cycle de vie de l'activité de l'intelligence d'affaires et 'analytique, dont celui de Gartner (voir figure 3).

Figure 3
Modèle de maturité BI



Source : Gartner (2008)

Comme l'indique la figure 3, le modèle de maturité de Gartner (2008) propose cinq niveaux de maturité à savoir : Inconscient (*Unaware*), Tactique (*Tactical*), Focalisé (*Focused*), Stratégique (*Strategic*) et Omniprésent (*Pervasive*). Il se concentre sur trois domaines clés que sont les personnes, les processus et les indicateurs/métriques ou la technologie. Le modèle de Gartner (2008) est généralement utilisé pour évaluer le niveau de maturité des entreprises, mais aussi la maturité des différents services (Oueslati, 2018). Par rapport à d'autres modèles, le modèle de Gartner (2008) offre une vue qui n'est pas uniquement centrée sur les aspects techniques, mais se concentre plutôt sur les aspects techniques des affaires de l'entreprise (Rajteric, 2010 ; cité par Oueslati, 2018). Cependant selon Oueslati (2018), les critères d'évaluation de la maturité ne sont pas réellement bien définis et la catégorisation repose principalement sur des niveaux de maturité des différents aspects évalués pris de manière individuelle. Toutefois, on remarque que chaque niveau de maturité offre une vision précise de ce qui est fait au niveau de l'organisation, ce qui sert de repère pour positionner la maturité d'une organisation quelconque, sur la base de sa pratique IA&A.

Au-delà de ce modèle de maturité BI, Gartner a conçu d'autres modèles de maturité qui comprennent les dimensions ou les axes de la maturité selon le

fonctionnement opérationnel de l'entreprise. Ces modèles renseignent sur la dimension de la gouvernance, des processus, de l'organisation, de la valeur d'affaires et bien d'autres. C'est l'exemple du modèle de maturité IAM (*Identity Access Management*), figure 4. Ce modèle qui porte sur la gestion d'accès est à propos dans cette rubrique. En effet, **l'accès aux systèmes et aux données est préliminaire à l'exploitation des données**. Le modèle de maturité IAM (figure 4) présente 5 niveaux de maturité qui débute au stade initial et s'achève au stade optimisé. Six dimensions sont évaluées grâce aux 5 niveaux de maturité. On voit d'après ce modèle qu'une organisation pourrait évaluer sa maturité selon la dimension de :

- **Gouvernance** : dans ce modèle, il est question de l'identité et la gestion des accès ; cependant, la dimension de gouvernance peut être évaluée pour n'importe quel processus dans l'organisation. Cette dimension établit un cadre de gestion dans l'organisation (Van Steenberge, 2017). Dans le modèle de maturité IAM, la dimension gouvernance permet d'évaluer où se situe l'organisation quant à la gestion de l'identité et des accès. Dans un autre modèle de maturité, elle permettrait d'évaluer la gestion de l'objet dont on évalue la maturité de cet objet dans l'entreprise. Il peut s'agir des données, de l'éthique, de la culture, etc.
- **Organisation** : cette dimension quant à elle, évalue comment l'objet dont on évalue la maturité est structuré et fonctionnel dans l'entreprise. Dépendamment du niveau de maturité, cette dimension peut être informelle. Les rôles basiques et la responsabilité décentralisée pour le niveau de maturité 1 (initial). Par contre, cette dimension dans une entreprise mature se manifeste par une intégration optimale avec l'entreprise et une centralisation des rôles.
- **Vision et stratégie** : cette dimension définit une image globale de ce que souhaite devenir l'entreprise au terme d'un horizon de planification qui a été préalablement retenue. Cette dimension guide l'entreprise dans la gestion du changement souhaité.
- **Processus** : cette dimension représente un ensemble d'activités corrélées ou interactives qui transforment les éléments d'entrée en éléments de sortie (norme ISO 9001:2000). Elle permet d'évaluer comment les activités sont

menées pour pouvoir transformer les données d'entrée en données de sortie. Comme l'indique la figure 4, à son stade initial, elle va être informelle, ponctuelle. Tandis qu'à son stade optimal (le plus grand niveau de maturité), elle va être optimisée après avoir été formalisée, intégrée et alignée avec les processus métiers de l'entreprise au stade 4.

- **Conception de l'architecture et de l'infrastructure** : cette dimension définit le squelette fonctionnel et matériel permettant de mener les activités quotidiennes. Selon le modèle architectural et de l'infrastructure d'une entreprise, on peut déterminer son niveau de maturité. En effet, l'architecture de l'objet (données, éthique, identité et accès, etc.) qu'on évalue sera alignée avec l'architecture de l'entreprise dans sa phase de maturité optimale, par contre, celle-ci va intégrer l'usage possible des certains outils à son stade initial.
- **Valeur d'affaires** : cette dimension permet de vérifier si l'entreprise mesure sa valeur d'affaires ou elle est plutôt rendue au stade l'optimisation et de l'amélioration de cette dernière ; c'est-à-dire qu'elle a déjà été mesurée et on cherche à l'optimiser.

Figure 4

Modèle de maturité IAM (Identity Access Management)

The Gartner IAM Program Maturity Model					
IAM Program Maturity Level	① Initial	② Developing	③ Defined	④ Managed	⑤ Optimized
Governance	Ad hoc, informal	Subsumed within InfoSec (and InfoSec governance structures)	IAM governance structure defined and accepted	IAM governance structure fulfilled and refined	IAM governance optimization
Organization	Informal, basic roles, responsibilities decentralized	Technical projects sponsored by BUs and CISO; informal inventory of IAM skills	IAM PMO established, IAM roles and training needs defined	IAM PMO active, RACI matrix defined; proactive skill development	Optimal integration with business; skills optimized
Vision and Strategy	Conceptual awareness at best	Certain business drivers identified; tactical priorities set	Business-aligned vision defined; strategic priorities set	IAM vision and strategy continually reviewed to track business strategy	Periodic optimization of vision and strategy
Processes	Ad hoc, informal	Semiformal BU-specific and target-specific processes	Formal processes defined, consistent across BUs and target systems	Formal processes integrated and refined; aligned with business processes	Process optimization
Architecture and Infrastructure Design	Possible use of target-specific productivity tools	Disjoint technical projects; technology redundancy likely	Discrete IAM architecture defined; rationalization and consolidation in hand	IAM architecture refined and aligned with EA	IAM architecture embedded within EA; optimization
Business Value	None measurable	Tactical efficiency and (maybe) effectiveness improvements; low direct value	Sustained, quantifiable improvements tied to GRC imperative; moderate direct value	Sustained, quantifiable contribution to all key business imperatives; high direct value	Business value optimization; transformational direct value
Legacy Program Maturity Level	Blissful Ignorance	Awareness	Corrective	Operational Excellence	

BU: business unit; IAM: identity and access management; PMO: program management office; EA: enterprise architecture; RACI: responsible-accountable-consulted-informed; GRC: governance, risk and compliance

Gartner

Source : Gartner (2010)

En résumé, les modèles de maturité BI et IAM ci-dessus présentés permettent de comprendre comment évaluer la maturité d'une organisation tant au niveau de sa pratique IA&A qu'au niveau de sa gestion organisationnelle de manière générale. Ces éléments seront utiles lors de l'analyse des résultats des experts. Ils permettront d'apprécier où se positionnent ceux-ci par rapport au modèle de maturité.

4. L'éthique, la morale et la déontologie

Les paragraphes précédents permettent de planter la base du décor de cette étude. En effet, les définitions des termes intelligences d'affaires, analytique et science de données aident à comprendre les éléments à aborder. Les modèles de maturité

complètent ce décor en fournissant les éléments qui permettent de positionner une organisation par rapport à ses pratiques IA&A et bien plus. Dans ce paragraphe, on va continuer à construire ce décor avec le concept de l'éthique, déontologie et morale. En effet, il est difficile de parler de concepts liés à l'exploitation éthique des données sans d'abord comprendre la nature même de l'éthique et de la déontologie. Ainsi dans ce paragraphe, on va premièrement définir le terme éthique, deuxièmement distinguer ce qui est éthique de ce qui est moral, de ce qui est légal et de la déontologie et troisièmement positionner l'éthique dans le contexte de l'intelligence d'affaires et de l'analytique.

a) L'éthique

Étymologiquement, le mot éthique vient du mot grec *ethos*. Celui-ci désigne les habitudes et les coutumes qui orientent l'agissement et qui sont elles-mêmes modifiées par celui-ci (Akamatsu, 2008; cité par François, 2010). Ludovic François reprend également la définition de l'éthique selon Aristote et Platon et déclare ceci : « l'éthique désigne le comportement visible qui repose sur des habitudes et des usages individuels et également sur des coutumes collectives et sociales » (François, 2010). Pour Ricoeur (1990), l'éthique est « la sagesse pratique » et vise « une vie accomplie sous le signe des actions estimées bonnes ». Quant à Luc Begin, il décrit l'éthique comme une réflexion sur les valeurs qui orientent et motivent les actions (Begin, 2015). Il ressort des définitions précédentes, que l'éthique se situe dans « l'agir » (François, 2010). Il est question en effet des actions qui sont posées ; du jugement qui accompagne ces actions. À ce titre, Richard Hackathorn, définit l'éthique comme « un jugement par les membres de la société sur ce qu'est un bon ou un mauvais comportement » (Hackathorn, 2003 ; Thomann et Wells, 2013). Pour ces auteurs, le comportement éthique, c'est donc de faire ce qu'on devrait faire. Hackathorn souligne que l'utilisation du mot « devrait » implique le jugement (Thomann et Wells, 2013). Il s'agit de la prise de décision de faire quelque chose de bon ou de mauvais ; il s'agit aussi de bon ou de mauvais comportement dans une société. C'est donc à ce titre que Frank Buytendijk décrivant l'éthique déclare qu'il s'agit de la conduite à suivre pour juger de ce qui est juste (Buytendijk, 2012) avec la prémisse que l'éthique est basée sur les concepts fondamentaux de soi, du bien et de

l'autre. L'éthique selon Buytendijk, considère ce qui est bon pour les autres aussi bien que ce qui est bon pour soi (Buytendijk, 2012 ; Thomann et Wells, 2013).

Les différentes définitions de l'éthique ci-dessus données ont en commun les concepts opposés du juste et du fallacieux, du bien et du mal ; ils soulignent ainsi la nature subjective du terme éthique. Juste et faux sont subjectifs, bien et mal sont également subjectifs. La subjectivité, selon Hackathorn, apporte l'ambiguïté et le besoin de jugement (Hackathorn, 2003). Or, Thomann et Wells pensent que le fondement des jugements éthiques doit être suffisamment solide pour rationaliser et concilier les diverses croyances ; un jugement fondé non pas sur des opinions, mais sur des valeurs, des principes et une pensée critique (Thomann et Wells, 2013).

On peut retenir de ce paragraphe que l'éthique est un ensemble de règles et d'orientations qui sont données aux individus par une source externe par exemple une entreprise ou une profession et qui permet de raisonner selon un but utilitaire, celui du bien-être du groupe.

b) L'éthique et la morale

Le paragraphe précédent définit le terme éthique selon son origine étymologique et selon les auteurs qui se sont déjà penchés sur la question. Il est question dans ce paragraphe de lever l'équivoque apparente entre l'éthique et la morale. En effet en quoi diffère l'éthique de la morale ? Existe-t-il une différence ? Les lignes suivantes permettent d'examiner la question.

Bien qu'ils soient étroitement liés, morale et éthique ne sont pas synonymes. Tous deux sont basés sur une philosophie du bien et du mal, mais avec des racines nettement différentes (Thomann et Wells, 2013). Pour Michele Bergadaà (2004), l'éthique et la morale se sont pendant plusieurs siècles inspirées des lois religieuses et politiques. Chacun pouvait interroger son dieu ou les dirigeants de sa cité pour savoir si son action était juste et bonne (Bergadaà, 2004). « Mais la morale semble néanmoins tombée en désuétude, car elle est souvent assimilée à un sentiment négatif du pouvoir (cf. « faire la morale ») » déclare Bergadaà (2004). Pourtant selon cette dernière, le sens originel de la morale n'est pas celui de l'interdit, mais de « liberté de l'homme qui décide en accord avec sa conscience ». Les entreprises américaines ont tenté de redonner une légitimité à

la morale en formalisant les règles d'une éthique pratique (Bergadaà, 2004), mais le glissement des définitions déjà vif dans la philosophie s'accroît dans le monde anglo-saxon qui confond généralement la morale, l'éthique et la déontologie sous la notion d'« *ethics* » (Isaac, 2000 ; cité par Bergadaà, 2004).

Pour Thomann et Wells (2013), l'éthique se rapporte à un ensemble de règles et d'orientations qui sont données aux individus par une source externe, telles qu'une entreprise ou une profession. Elle conduit à raisonner selon un but utilitaire, celui du bien-être du groupe (Bergadaà, 2004). La morale quant à elle, est liée selon Thomann et Wells (2013) aux croyances et aux principes personnels d'une personne au sujet du bien et du mal.

En effet, l'origine étymologique de la morale est le mot latin *mores*, qui signifie le « caractère ». Kant a le mérite de considérer la morale comme le fruit d'un être qui raisonne (Bergadaà, 2004). Pour cet auteur, la morale ne découle pas d'une loi externe imposée à l'individu, mais elle possède un caractère universel. C'est à ce titre qu'il déclare : « *Agis toujours de telle sorte que la maxime de ton action puisse être érigée en loi universelle* ».

Il ressort de ce qui précède que l'éthique autant que la morale sont fondées sur la notion du bien et du mal. **Cependant l'éthique diffère de la morale en ceci que cette dernière ne découle pas d'une loi externe imposée à l'individu comme c'est le cas de l'éthique, mais est liée aux croyances et aux principes personnels d'une personne au sujet du bien et du mal.**

Dans le paragraphe suivant, il est question d'examiner ce qui est éthique par rapport à ce qui est légal.

c) L'éthique et la légalité

Après avoir présenté en quoi l'éthique diffère de la morale dans la section précédente, il est question dans cette section de voir en quoi elle diffère de la loi.

En effet, ce qui est légal et ce qui est éthique sont des concepts connexes et qui se chevauchent, mais ce sont des choses discrètes et séparées (Thomann et Wells, 2013). Pour ces auteurs, ce qui est légal se réfère simplement aux actions qui sont conformes aux lois ou aux statuts. Être éthique par contre, signifie faire ce qui est juste ou mieux

faire la bonne chose. Il y a plus d'ambiguïté dans les questions éthiques que dans les questions juridiques si on s'en tient au caractère subjectif de l'éthique (de la notion du bien et du mal) tel que décrit dans les sections précédentes. En effet, **la loi est prescriptive et requiert de s'y conformer, sinon, des sanctions préalablement établies s'en suivent**. Généralement, ce qui est perçu comme étant contraire à l'éthique est très souvent illégal. Cependant, l'inverse, c'est à dire la perception que ce qui est légal doit aussi être éthique, est moins courant. Ce qui est légal peut encore être perçu comme contraire à l'éthique, et ce qui est illégal peut parfois être perçu comme éthique. La perception peut façonner l'éthique, mais ne façonne pas le droit. La conformité juridique est un mandat pour les entreprises et les sociétés. La conduite éthique est un choix. Pour donner un exemple concret : Acme Toyz est un service de jeu en ligne. Les utilisateurs doivent s'inscrire pour pouvoir jouer à ses jeux, qui peuvent être joués par n'importe qui, y compris les enfants de moins de 18 ans. Pour Acme Toyz, la question éthique est de savoir s'il est acceptable (« bon ») de recueillir des renseignements sur le nom et l'adresse des enfants sans le consentement de leurs parents. La question morale est de savoir si les enfants de moins de 18 ans devraient avoir accès à des jeux violents du tout (avec ou sans le consentement des parents). La question juridique est de savoir si l'entreprise est autorisée à collecter des données d'utilisateurs mineurs et si elle le fait, quelles sont les sanctions prévues si les parents intentent une action en justice contre l'entreprise.

Il convient de retenir de cette section **que l'éthique et la notion de légalité sont deux termes qui se chevauchent, mais qui sont discrets et distincts. Ce qui est légal se réfère aux actions qui sont conformes aux lois ou aux statuts ; être éthique par contre, signifie faire ce qui est juste ou mieux, faire la bonne chose en se basant sur les principes prescrit par une entité ayant autorité comme une entreprise ou un regroupement notamment un ordre professionnel.**

d) L'éthique et la déontologie

Le paragraphe précédent montre succinctement la différence entre l'éthique et la notion de légalité. Le but est de permettre une compréhension de chacun des termes afin

de mieux entrevoir l'utilisation de chacun d'entre eux. Qu'en est-il de l'éthique par rapport à la déontologie ? Les lignes suivantes traitent la question posée.

Afin de normaliser les décisions publiques, Bentham a introduit pour la première fois le concept « **déontologie** » (Bergadaà, 2004). L'auteur jugeait le terme plus expressif que celui de l'éthique. Pour lui, **la déontologie est assimilée à un mouvement ou à une action qui accompagne l'exercice de la fonction**. Luc Begin (2015) appuie cette position et mentionne pour sa part que l'éthique diffère de la déontologie en ceci que cette dernière renvoie aux obligations que des personnes sont tenues de respecter dans leur travail (Begin, 2015). Ces personnes pourraient être des travailleurs d'une même profession par exemple des enseignants ou des ingénieurs ; des personnes au service d'un même employeur, comme les employés de l'université de Sherbrooke ; des gens exerçant des fonctions professionnelles semblables comme des travailleurs d'un même secteur notamment le milieu des affaires.

Étymologiquement, la « déontologie » vient des mots grecs *deon* et *logos* soit le *devoir* et le *discours*. D'après le dictionnaire Larousse, la déontologie va se définir comme « *l'ensemble des règles et des devoirs qui régissent une profession, la conduite de ceux qui l'exercent, les rapports entre ceux-ci et leurs clients ou le public.* »

La déontologie d'après Bergadaà, s'est « particulièrement » développée dans les entreprises et services où elle réduit le risque perçu par les utilisateurs et les professionnels. Si la déontologie indique le sens des actions, l'éthique recommande d'agir d'une manière ou d'une autre en fonction du cadre d'action auquel on se réfère (Bergadaà, 2004).

Comme mentionné dans le paragraphe éthique et moral du présent document, les termes déontologie, morale et éthique ont un sens moins clair dans le monde anglo-saxon. Ces trois sont tous confondus sous le terme sous la notion d'« *ethics* » (Isaac, 2000 ; cité par Bergadaà, 2004).

Cependant, il faut noter que malgré la diversité des activités, des organisations et des situations, les actions d'éthiques peuvent être regroupées en cinq domaines dont quatre correspondent à des valeurs ou plutôt à des constellations de valeurs, que l'on retrouve dans la plupart des référentiels éthiques d'entreprise, et le cinquième est plutôt

une obligation réglementaire qui s'impose à l'entreprise tant sur le plan national qu'international (Begin, 2015) :

- Le respect des personnes ;
- L'intégrité ;
- Les relations avec les acteurs commerciaux (clients-consommateurs / fournisseurs-prestataires/ concurrents) ;
- Le respect de l'environnement ;
- Le respect des lois et réglementations externes ou internes.

En résumé, on peut retenir de ce paragraphe **que l'éthique diffère de la déontologie en ceci que cette dernière renvoie aux obligations que des personnes sont tenues de respecter dans leur travail ; elle est assimilée à un mouvement ou à une action qui accompagne l'exercice de la fonction.**

5. L'éthique en intelligence d'affaires

L'IA&A est chargée d'ambiguïtés et de tensions entre différentes perspectives du bien et du mal (Thomann et Wells, 2013). Dans les paragraphes précédents, on va vu que l'éthique, la morale et la déontologie font intervenir plusieurs concepts. Dans ce contexte, on utilise les éléments présentés dans le but d'outiller les professionnels en exploitation des données à faire face aux ambiguïtés et tensions entre diverses perspectives du bien et du mal tel qu'évoqué par Thomann et Wells (2013). L'idée est d'adopter une approche pragmatique à leur niveau et non de faire une étude sur le concept éthique, morale ou déontologie. On en parle ici dans le but de pouvoir mieux formaliser la présente recherche et mieux comprendre le cadre dans lequel les actions des professionnels en exploitation des données pourraient se positionner, tout en évitant de faire une étude approfondie de ces concepts, car ce n'est pas le mandat de la présente recherche.

« La conduite éthique rapporte des récompenses et un comportement contraire à l'éthique comporte des risques. » (Thomann et Wells, 2013), c'est conscient de cette

vérité que la présente section traite de la question d'éthique en intelligence d'affaires et l'analytique.

En effet, la question de savoir si la conduite d'activités liées à l'intelligence d'affaires et l'analytique est éthique ou non comporte en réalité deux volets (Schultz et al., 1994; Demilia et al., 2012). Premièrement, il y a la question générale de savoir si la collecte et l'utilisation des données constituent ou non une pratique éthique. Pour ces auteurs, si on arrive à la conclusion que cette « fin » est acceptable, alors il faut se pencher sur la question des moyens. Cette deuxième question porte sur les méthodes particulières de collecte et d'utilisation des données qui sont permises et celles qui ne le sont pas. Il convient donc d'en retenir que les comportements éthiques en intelligence d'affaires et analytique vont être donc à la fois au niveau de la collecte et de l'utilisation des données qu'au niveau des moyens employés.

À ce sujet, Thomann et Wells (2013) mentionnent que les questions éthiques se posent le plus souvent lors de la collecte de données, particulièrement les données sur les clients et les concurrents. Leur affirmation est sans doute due au fait que les données sur les clients et les consommateurs sont peut-être les plus sensibles sur le plan éthique parce qu'elles sont les plus susceptibles de causer du tort sous des formes qui vont des pratiques de marketing trompeuses au vol d'identité. Si ces données sont parfois internes à l'entreprise, il faut dire que des pratiques bien que légales, mais peu ou pas du tout éthiques sont souvent employées pour la collecte. Il s'agit des techniques comme : a) **l'information trompeuse** (en anglais, « *misrepresentation* ») (François, 2010; Thomann et Wells, 2013.), qui peut consister à se faire passer pour quelqu'un afin d'obtenir des informations comme un étudiant, un client, etc. ou encore utiliser des personnes ne mentant pas sur leur fonction ni sur leur identité, mais sur l'objet de l'entretien ; par exemple un journaliste travaillant pour un cabinet d'intelligence d'affaires et l'analytique. b) La « **manipulation** » qui consiste à amener les consommateurs à révéler des renseignements personnels (Thomann et Wells, 2013). Elle est surtout répandue dans l'économie numérique. De nombreuses permissions d'applications pour smartphones incluent l'accès à des informations personnelles dont ils n'ont pas vraiment besoin. À cela s'ajoutent des applications mobiles sensibles à l'emplacement. La réalité est que l'utilisateur ne sait pas quelle est la quantité de données recueillies ; combien de temps est-elle conservée et comment s'en sert-on ?

D'autres techniques moins trompeuses sont également utilisées, notamment l'achat de données auprès de vendeurs de données, l'achat de données auprès de consommateurs (offrir un cadeau ou une remise en échange de la divulgation de données personnelles), les programmes de fidélité, etc. (Thomann et Wells, 2013.) Pour ces auteurs, chaque méthode a des répercussions sur les questions relatives aux principes d'éthique des données.

Thomann et Wells (2013) ne se limitent pas aux considérations éthiques au niveau de la collecte des données ainsi que des moyens utilisés à cette fin. Parlant de l'utilisation des données, ils vont dire « les données elles-mêmes ne sont ni bonnes ni mauvaises en soi. Ce sont les utilisations que nous faisons des données et l'intention de ces utilisations qui exigent un jugement éthique entre le bien et le mal » (Thomann et Wells, 2013). En effet, la pratique de l'intelligence d'affaires et l'analytique repose aussi sur l'utilisation des données dans le but de soutenir les processus décisionnels. Pour Thomann et Wells (2013), une bonne décision commence par une bonne information, laquelle est caractérisée par deux éléments clés, dont l'exactitude et l'exhaustivité. L'information présentée pour soutenir une décision peut donc être sujette à un choix éthique notamment au sujet de ce qu'il faut présenter ou non. Les professionnels en exploitation de données ont la capacité de biaiser les décisions : par exemple, ils peuvent amplifier certaines tendances tout en cachant d'autres. Les données visuelles créant de nombreuses occasions d'ajouter des biais, comme en témoigne le livre populaire *How To Lie With Statistics* (Huff, 1954). On pourrait donc se poser la question de savoir quand est-il éthique d'introduire des préjugés dans les données et quand est-ce contraire à l'éthique ? Dans le paragraphe précédent, on a vu que plusieurs techniques sont utilisées pour collecter les données. Une des questions posées à ce stade était alors de savoir ce qui est fait de ces données. Thomann et Wells (2013) fournissent une partie de la réponse. En effet, pour ces derniers, parfois, les données sont utilisées pour améliorer l'expérience client ; parfois, elles sont utilisées pour mener des campagnes de marketing très ciblées ; parfois, elles deviennent une source de revenus supplémentaire lorsque les données sont revendues ; parfois, elles créent un risque de vol d'identité (Thomann et Wells, 2013). En effet, le profilage des consommateurs est une application de plus en plus courante des données. Comme mentionné précédemment, souvent sans le savoir, les consommateurs fournissent des

données personnelles dans leurs activités quotidiennes. Les transactions par carte de crédit, la navigation sur le site Web, les achats en ligne, l'utilisation des guichets automatiques, les activités de médias sociaux et l'utilisation du téléphone mobile. Tout ceci génère des données qui sont recueillies, agrégées, syndiquées, vendues et utilisées pour mieux comprendre les consommateurs et leurs comportements. Chacun de ces scénarios a des répercussions sur les principes d'éthique des données. Il devient donc difficile d'établir la limite entre le moment où les données sont utilisées pour améliorer l'expérience client et persuader ce dernier et lorsque l'utilisation des données devient de la manipulation.

Il faut dire que plusieurs autres utilisations des données sont possibles, notamment la production de rapports, l'analyse, l'exploitation, la découverte, la détection des fraudes, et plus. Cependant, chacune d'elles a des répercussions sur les questions relatives aux principes d'éthique des données. Il convient donc d'être intentionnel et d'outiller les professionnels en IA&A d'un code pour les aider dans leurs pratiques quotidiennes.

En somme, la présente section traite de l'éthique en intelligence d'affaires et analytique. Elle rappelle que les considérations éthiques existent tant au niveau de la collecte et l'utilisation des données qu'au niveau des moyens utilisés. Parlant de la collecte, il a été mentionné que les données des clients et les concurrents sont celles qui posent le plus souvent les problèmes éthiques. Les moyens de collectes sont eux aussi sujets à des problématiques éthiques, car il s'agit parfois d'information trompeuse ou de manipulation. Au sujet de l'utilisation des données, les enjeux éthiques sont également mentionnés, notamment l'introduction des biais, la mauvaise présentation des données où on fait par exemple mentir les chiffres par la façon de rendre l'information, l'utilisation abusive des données collectées à des fins non avantageuses au client, mais plutôt à l'entreprise. La liste des problèmes éthiques peut s'allonger considérant les autres utilisations faites des données, dont la production des rapports, l'analyse, etc.

Dans la section suivante, quelques enjeux éthiques ainsi que les éléments qui y sont liés seront présentés.

6. Quelques enjeux éthiques en intelligence d'affaires

Dans la section précédente, on montre que des questions éthiques existent dans la pratique des activités de l'intelligence d'affaires et l'analytique. Il est question dans cette section d'en énumérer quelques-uns et voir comment ils sont abordés dans la littérature scientifique.

En s'appuyant sur ce qui précède, on peut dire sans risque de se tromper qu'il existe plusieurs enjeux éthiques en intelligence d'affaires et analytique parmi lesquels :

- **La protection de la vie privée** : Demilia et al. (2012) mentionnent que la nature de l'information collectée et la mesure dans laquelle cette information est recueillie ont toutes deux des ramifications éthiques. Ces auteurs ressortent la question relative à la protection de la vie privée comme un problème éthique directement attachée à l'activité de la collecte des données. Thomann et Wells (2013) soutiennent cette position et s'expriment au travers des questions comme : où s'obtiennent les données ? Qui collecte les données ? Comment les données sont-elles obtenues ? Leurs opinions sur la vie privée des clients sont parfois violées par des activités liées au métier de l'intelligence d'affaires et l'analytique. En effet, dans la section précédente, il a été établi que des applications informatiques ont accès aux données dont elles n'ont pas besoin, contribuant ainsi à la violation de la vie privée et par la même occasion soulevant un second enjeu qui est celui de l'accessibilité à l'information. Les questions posées par Thomann et Wells (2013) laissent entrevoir que les utilisateurs ne sont pas toujours au fait que leurs données personnelles sont collectées ou lorsqu'ils le sont, ils ignorent la nature des données collectées.
- **L'accessibilité à l'information** : comme mentionné dans le paragraphe précédent, l'enjeu de l'accessibilité est lié à celui de la protection de la vie personnelle dans une certaine mesure. En effet, par l'activité de collecte, les objets connectés ont accès à des données incluant des données personnelles. Aussi les usagers laissent des informations dont ils n'ont pas nécessairement la possibilité de décider de leur finalité lorsqu'ils utilisent les plates formes informatiques (réseaux sociaux, site de commerce en ligne, site des fournisseurs de services). En effet, une fois des données laissées ou

collectées sur ces plates formes, les utilisateurs n'y ont plus toujours accès. Cependant, des personnes pas toujours bien intentionnées y ont accès et les utilisent parfois de manière à causer des dommages aux vrais propriétaires des données. Ce problème a entre autres conduit la Commission européenne à introduire dans le RGPD (Règlement Général sur La Protection des Données), une clause permettant aux usagers d'avoir accès à leurs données personnelles collectées s'ils le souhaitent. Il s'agit de l'article 15, considérants 63 et 64 (RGPD, 2018).

La littérature scientifique soutient que les enjeux éthiques aboutissent ou sont directement liés à de vrais problèmes parmi lesquels :

- **Consentement éclairé** : notamment au niveau de la collecte de données. Il est question, ici, de savoir si le sujet des données connaît que des données à son sujet sont recueillies, et si son autorisation de l'activité de collecte de données est requise (Thomann et Wells, 2013 ; Demilia et al., 2012).
- **Anonymat** : Thomann et Wells (2013) lancent la réflexion à travers ces deux questions: tous les renseignements personnels d'identification devraient-ils être éliminés des données ? Les données devraient-elles être recueillies uniquement sous forme d'agrégats de sorte que les personnes ne puissent être identifiées ? Il ressort de ces questions que l'anonymat est le fait de ne pas permettre l'identification de la personne dont on recueille les données. Cependant, McBride (2014) l'aborde via la manière dont l'identité est affectée par le système. Pour lui, la question qui se pose est : comment le système d'information affecte-t-il l'identité et la finalité de l'utilisateur ? On pourrait entrevoir ici l'idée de croisement de certaines informations stockées dans les systèmes informatiques et à la base dépouillée de toute donnée personnelle avec d'autres informations et arrivées à briser l'anonymat.
- **Confidentialité** : Les sources et les fournisseurs de données devraient-ils être protégés contre la divulgation ? Se questionnent Thomann et Wells (2013). La question laisse sous-entendre que la confidentialité consiste à protéger de la divulgation de la donnée.
- **Sécurité** : Thomann et Wells (2013) abordent la notion à travers la question : dans quelle mesure les données devraient-elles être protégées contre l'intrusion,

la corruption et l'accès non autorisé ? On peut voir une fois de plus par le questionnement de ces auteurs que la sécurité consiste à protéger les données contre l'intrusion, la corruption et l'accès non autorisé. Pour Paulsen et Coulson (2011), il s'agit de protéger l'intégrité, la disponibilité et la confidentialité des données et des systèmes, qui peuvent être essentielles au maintien des activités d'une organisation. Même si ces deux définitions semblent a priori différentes, on se rend compte que ce n'est pas tout à fait le cas, mais qu'elles se complètent. En effet, dans la définition de Thomann et Wells (2013), il est question de la protection contre la corruption et l'intrusion. Ces deux éléments s'ils sont respectés viennent garantir l'intégrité dont mentionnent Paulsen et Coulson. La protection des accès non autorisés intégrée dans la définition de Thomann et Wells vient garantir la confidentialité dont il est question dans la définition de Paulsen et Coulson. Et la disponibilité évoquée par Paulsen et Coulson est un élément complémentaire. L'autre élément complémentaire est le fait que ces derniers ne limitent pas leur définition à la protection des données, mais l'étendent à la protection des systèmes.

- **Transparence** : McBride (2014), aborde la réflexion sur ce concept par la question suivante : la dérivation et l'utilisation de l'information sont-elles claires pour les utilisateurs ? On voit au travers de la question de cet auteur que la transparence doit consister à informer clairement quelle est l'utilisation de la donnée et quelles sont les dérivations de la donnée. Alors que McBride adresse le concept en se questionnant, Turilli et Floridi (2009) donnent leur définition du concept comme consistant à rendre visible l'information de sorte que son contexte soit facilement accessible, que sa provenance et son origine soient divulguées et que les processus humains et opérationnels qu'elle soutient soient clairement expliqués et justifiés. Comme pour le cas de la sécurité, on voit ici des définitions complémentaires de la part des auteurs pour le concept de transparence.
- **L'exactitude** : il s'agit pour Ming et al. (2015) de la légitimité, de la précision et de la pureté avec lesquelles l'information est réalisée.
- **Secret** : Pour Ming et al. (2015), le terme fait référence aux détails relatifs à la propriété, à l'équité de l'échange et aux caractéristiques d'accès à la donnée.

McBride (2014) quant à lui, affirme qu'il est désormais impossible de garantir le caractère secret en raison de l'accessibilité des données pour tous.

- **Précision** : Thomann et Wells (2013) reviennent en se demandant : Quel est le niveau d'exactitude et de justesse exigé des données ? Au travers de la question, les auteurs laissent entrevoir que la précision représente le niveau d'exactitude et de justesse exigé des données.
- **Honnêteté** : Dans quelle mesure l'entreprise devrait-elle paraître directe et visible au sujet des pratiques de collecte de données ? La question est encore posée par Thomann et Wells (2013) afin de susciter la réflexion sur l'honnête de la compagnie qui manipule les données. On peut remarquer ici que les auteurs ne parlent pas de la donnée en soi, mais plutôt de l'entreprise qui collecte la donnée. Pour eux, l'honnêteté doit consister à exposer de manière directe et visible ses pratiques de collecte de données.
- **Propriété** : Ming et al. (2015) disent qu'il s'agit de l'information sur soi qu'une personne est prête ou forcée à abandonner. En considérant ce terme sous l'angle de McBride (2014), on ne devrait pas se préoccuper de la propriété des données en soi si celles-ci sont disponibles sur le web. En effet, il justifie sa position par le fait que l'un des principes fondateurs de l'Internet était que l'accès et le contenu seraient gratuits. Pour lui, Berners-Lee (fondateur du www) ne cherchait pas à obtenir des droits de propriété intellectuelle. Celle-ci est appuyée par Maxwell (2004) qui affirme qu'il est discutable de qualifier l'information de « propriété ».
- **Responsabilité** : Qui est responsable et à quel niveau pour l'utilisation et la mauvaise utilisation des données ? questionnent Thomann et Wells (2013). En posant la question, Thomann et Wells laissent comprendre que quelqu'un doit répondre de l'utilisation des données, que cette utilisation soit bonne ou pas.

En résumé, cette section traite des enjeux éthiques en intelligence d'affaires et analytique. Il a été évoqué quelques-uns, dont la protection de la vie privée et l'accessibilité aux données personnelles, bien que ces deux éléments ne représentent pas la totalité des enjeux si on s'en tient au contenu du paragraphe 6 sur l'éthique en intelligence d'affaires. Il est également question dans cette section des problèmes relatifs à la collecte et à l'exploitation des données, dont l'anonymat, la confidentialité,

le caractère secret et plus. Ici les concepts ont été définis selon la littérature existante afin de permettre une compréhension de ce que peut représenter chacun d'entre eux.

Dans le paragraphe suivant, il est question de présenter quelques ancrages déontologiques présents dans le monde des affaires. Ici, on va premièrement décrire comment les autres disciplines d'affaires sont organisées à travers le monde et deuxièmement, on verra ce qui en est exactement au niveau de l'intelligence d'affaires en ce moment.

7. Ancrages déontologiques

Plusieurs disciplines composent le monde des affaires, dont la comptabilité, le marketing, les ressources humaines, l'intelligence d'affaires et l'analytique, etc. Certaines sont régies par des codes de déontologie balisant le fonctionnement de la discipline. Il est question dans cette section de présenter premièrement des exemples de code d'éthiques et de déontologie pour ces disciplines, deuxièmement de voir les points de convergences et de divergence par rapport à l'IA&A et enfin de voir les initiatives menées pour l'encadrement de la pratique éthique de la discipline de l'IA&A.

a) Codes d'éthique et de déontologie

❖ Cas des ressources humaines

Au pays, la discipline des ressources humaines est depuis 1961, organisée autour d'une association qui a tour à tour évolué vers la société des conseillers en relations industrielles en 1963 pour devenir dix ans plus tard, corporation professionnelle et par la suite ordre professionnel des conseillers en relations industrielles du Québec (CRHA, 2019). En, 1977, elle approuve par un règlement un code de déontologie (CRHA, 2019). On constate que 15 années seulement après son organisation, la discipline des ressources humaines est structurée et fonctionne sous la balise d'un code de déontologie qui **régit les activités de ses membres et aide à augmenter la confiance des clients et partenaires d'affaires**. Son code de déontologie couvre les questions éthiques qui surviennent autour de la relation entre l'employeur et l'employé notamment les devoirs et les droits de chaque partie envers l'autre dans la relation. On y retrouve par exemple

des éléments liés à la discrimination sur la base de l'âge, du genre, de l'ethnie, de l'orientation sexuelle, des convictions religieuses, du poids ou de l'apparence physique. Par cet élément, l'ordre des ressources humaines met l'accent sur l'égalité et vient rencontrer un des principes mêmes de droit fondamental de l'individu. On retrouve également dans le code de déontologie des ressources humaines, des éléments liés au respect de la vie privée des employés, de l'équité du contrat de travail, de l'hygiène et sécurité au travail et même de la dénonciation par un (ancien) employé de comportements répréhensibles cachés par son employeur. Ces éléments sont de nature à baliser le travail du professionnel en le protégeant, mais aussi en protégeant le public (travailleurs, dirigeants, gestionnaires et organisations) qui bénéficie de ses services (CRHA, 2019). À titre d'exemple, l'article 51 qui régit le secret quant aux renseignements de nature confidentielle qui viennent à la connaissance du professionnel des ressources humaines dans l'exercice de sa profession stipule que ce dernier doit :

- S'abstenir de faire usage de tels renseignements au préjudice du client ou en vue d'obtenir, directement ou indirectement, un avantage pour lui-même ou pour autrui;
- Prendre les mesures nécessaires pour que ses collaborateurs et les personnes qu'il a sous son autorité ou sa supervision ne divulguent pas ou ne se servent pas de tels renseignements qui viennent à leur connaissance dans l'exercice de leurs fonctions;
- Éviter de tenir ou de participer à des conversations indiscrètes au sujet d'un client et des services qui lui sont rendus.

❖ **Cas de la profession comptable**

La discipline comptable tout comme la discipline des ressources humaines est organisée autour d'un ordre professionnel. Au Canada, chaque province avait un ordre comptable pour les professionnels de son territoire. Les premiers ayant été constitués en 1879 avec l'ordre des comptables du Québec, et de l'Ontario. En 1925, le premier comité d'éthique professionnel est créé et travaille à l'élaboration des lignes de conduite nationale en matière d'éthique dont l'adoption sera faite en 1947. Cette loi adoptée en 1947 évolue et s'adapte au contexte de la profession (CPA, 2019). Son code de

déontologie a pour principal but : **encadrer la pratique des professionnels**. Il prévoit qu'un professionnel comptable doit rendre des services professionnels conformes aux normes techniques et professionnelles applicables. Qu'il est tenu de suivre avec diligence et compétence les instructions de son client ou de son employeur dans la mesure où elles sont compatibles avec les exigences d'intégrité et d'objectivité. Le code de déontologie comptable s'attaque aux questions comme celles de la rémunération des dirigeants ou celles de l'utilisation illicite à des fins privées des ressources de l'entreprise. Il s'attaque aussi aux questions en rapport avec le comportement moral des professionnels notamment en ce qui a trait à l'acceptation des pots de vins ou de conflit d'intérêts.

❖ **Cas du marketing**

Bien que la discipline ne soit pas organisée sous un ordre national ou provincial, les professionnels du marketing, organisé autour de l'association canadienne de marketing se sont dotés d'un code d'éthique qui auto réglemente la pratique de leur métier au quotidien. Ce code intègre les éléments régissant les pratiques en matière de renseignements personnels, de la limite d'une campagne, la véracité notamment au sujet de la présentation d'un agent marketing à un consommateur ou à une entreprise et la clarté quant à la communication faite par le professionnel avec un client (CMA, 2019). D'autres éléments en rapport avec la relation entre le professionnel et le public y sont mentionnés afin de **favoriser une pratique intègre de la discipline**.

❖ **Cas de l'intelligence d'affaires et l'analytique**

Contrairement aux autres disciplines du domaine d'affaires, l'IA&A ne fonctionne pas sous une dénomination de l'« ordre » ou d'« association ». Elle n'a donc pas un code de référence éthique ou déontologique régissant sa pratique. Cependant à l'échelle internationale, les professionnels de l'intelligence économique en France regroupés sous la Fédération des Professionnels de l'intelligence économique (FEPIE) s'engagent à respecter un code d'éthique qui intègre des éléments visant à préserver les intérêts fondamentaux de la France, à n'avoir recours qu'à des moyens légaux quel que soit le lieu d'application de leur activité, à intégrer une clause de confidentialité dans

tout contrat entre parties, à éviter les conflits d'intérêts, à fournir des informations accessibles par des moyens légaux (FEPIE, 2019).

Toujours à l'international, La *Strategic and Competitive Intelligence Professionals* (SCIP) qui est un organisme agissant à une échelle mondiale pour les professionnels engagés dans l'intelligence stratégique, intégrée et compétitive et les disciplines influençant les décisions organisationnelles connexes propose quant à lui un code de déontologie qui intègre des **éléments liés aux comportements de ses membres** notamment le respect des lois applicables nationales et internationales, la divulgation avec exactitude de toutes les informations pertinentes à l'identification du membre lors d'une entrevue par exemple ; d'éviter les conflits d'intérêts, fournir des recommandations honnêtes et réalistes dans le cadre de leur fonction (SCIP, 2019)

Une fois de plus à l'échelle internationale, le groupe européen pour l'éthique en science et technologie a commandité un travail portant sur l'éthique en intelligence artificielle particulièrement sur la robotique et les systèmes autonomes. Ce travail dont le manuscrit a été déposé en mars 2018 fait des constats sur les avancées technologiques surtout en matière d'apprentissage machine, de systèmes automatisés... ensuite aborde des questions d'ordres moraux.

À l'échelle nationale, on remarque, concernant l'intelligence artificielle et les données massives que la littérature s'intéresse aussi à l'éthique de ces pratiques. Et force est de constater que les organisations, aussi bien les entreprises, mènent des réflexions autour de la question afin de trouver la bonne façon de faire les choses. À l'échelle nationale par exemple, le gouvernement fédéral annonce sur son site web l'utilisation de l'intelligence artificielle dans ses décisions tout en évoquant clairement quelques principes directeurs dont la transparence, la clarté du processus décisionnel accompagnant sa démarche (gouvernement fédéral du Canada, 2019). S'il ne s'agit pas encore d'un code d'éthique dédié à l'intelligence artificielle pour les professionnels du gouvernement fédéral, il est cependant impossible de nier la volonté d'adopter la démarche éthique et d'intégrer la question dans leur quotidien. Plusieurs autres organismes à l'échelle nationale sont préoccupés par la question et certains d'entre eux offrent des formations sur l'éthique à leurs technologues. En effet, une étude menée en

septembre 2018 indiquait que 63% des entreprises canadiennes du domaine de l'intelligence artificielle offraient une formation d'éthiques à leur technologue tandis que 73% organisent un examen hebdomadaire pour surveiller les pratiques de l'intelligence artificielle (avant-première, 2019). Les initiatives comme celle de Datascience Montréal ou de DataDriven Montréal sont aussi à souligner et visent à conscientiser les membres et professionnels sur les enjeux éthiques liés à leurs professions et les comportements à adopter face à cela. (DSDTMTL, 2019; DataDrivenMTL, 2019).

Conclusion

Cette section traite la revue de littérature. On y a défini les concepts d'intelligence d'affaires, de sciences de données, d'éthique et brossé le portrait d'un scientifique de données puis on a présenté quelques enjeux éthiques selon la littérature existante. On y a également abordé la différence entre l'éthique et la morale, l'éthique et la déontologie et enfin l'éthique et le légal. Ce point a permis de réaliser que même si les concepts sont proches, ils ont quelques spécificités qui permettent de les différencier ; toutefois, on a vu qu'en Amérique du Nord, on ne tient pas compte de ces spécificités et les concepts éthique, moral et déontologie sont combinés et utilisés sous le terme « *ethics* ». Il a également été question de présenter le modèle de maturité en intelligence d'affaires afin comprendre et de situer la maturité d'une entreprise dans le domaine de l'intelligence d'affaires, en fonction de ses pratiques en la matière. Pour terminer, on a présenté quelques exemples de code d'éthiques des disciplines d'affaires ainsi que les initiatives qui se font dans le domaine de l'intelligence d'affaires. Cette section permet ainsi de définir le cadre théorique de la présente étude. Dans la section suivante, il est question de définir la problématique de recherche. En effet, après avoir planté le décor du cadre théorique, on a les bases qui permettent de comprendre la problématique et de la définir. La section suivante y est consacrée.

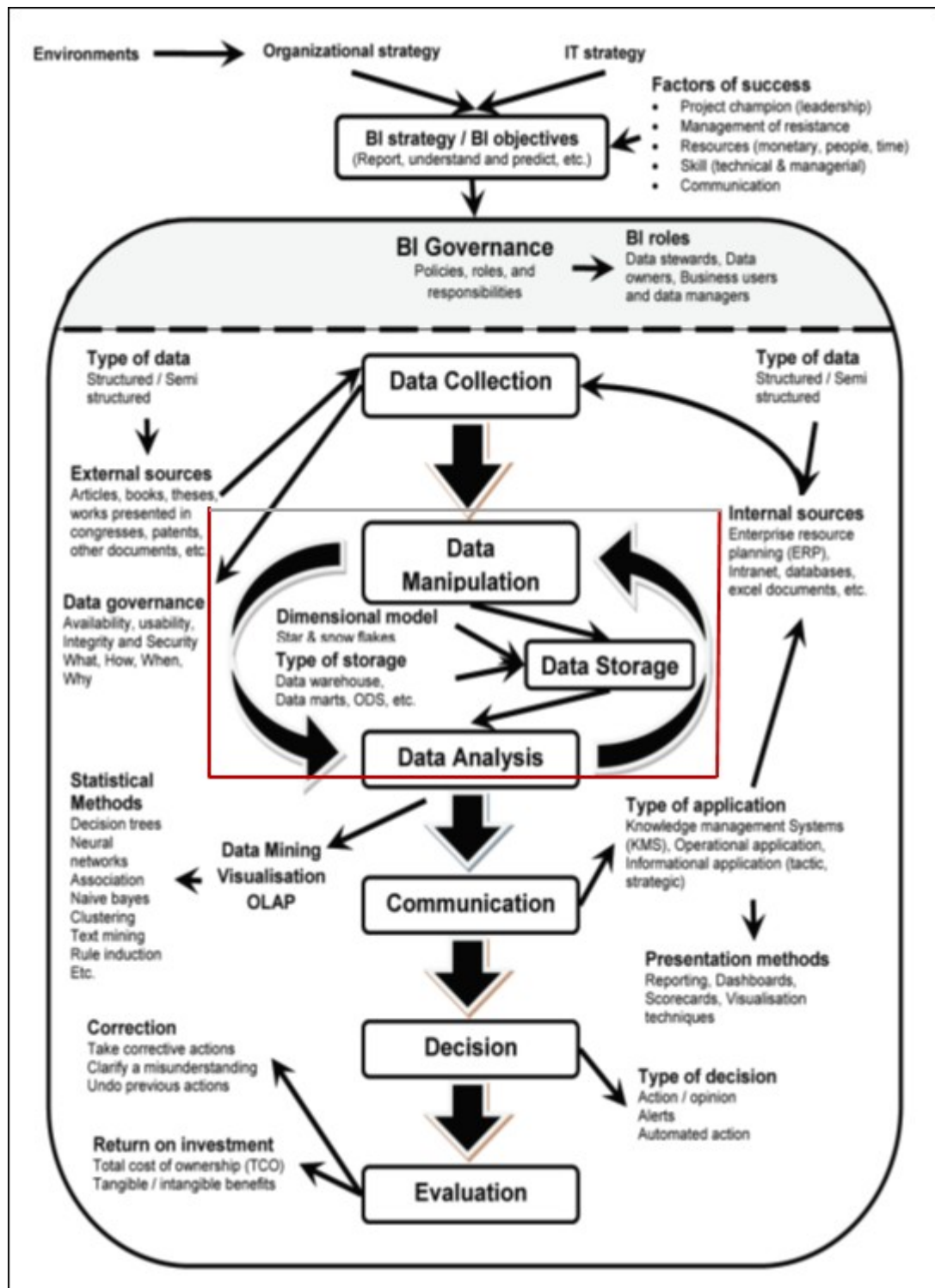
CHAPITRE 3 : PROBLÉMATIQUE

Cette section présente la problématique de recherche. Dans un premier temps, on fait état de la pertinence et de la nouveauté du sujet choisi pour cette recherche. Ensuite, on formalise la problématique de recherche. Enfin, on énonce les objectifs de la recherche.

1. Mise en contexte de la recherche

La collecte et l'exploitation des données en contexte d'affaires sont devenues des activités d'une importance névralgique dans les organisations et les entreprises de toute taille. Le virage numérique des organisations a permis l'émergence de nouveaux métiers comme les scientifiques de données, les ingénieurs des données, les spécialistes de l'analytique, les analystes de données, intendants de données, les intégrateurs de données. Ces rôles, comme plusieurs autres, sont placés sous l'égide de l'intelligence d'affaires et l'analytique. Des programmes académiques de formation technique et d'affaires ont été élaborés pour aider les nouveaux professionnels à acquérir les compétences nécessaires pour accomplir leur travail. Ce domaine d'affaires vise la gestion et l'exploitation des données pour la prise de décision concernant les concurrents, les fournisseurs, les clients, les problèmes financiers, les enjeux stratégiques, les produits et les services à tous les niveaux de l'organisation. Il aide aussi à rendre compte de la performance de l'entreprise, et à découvrir de nouvelles occasions d'affaires. Elle opère selon un processus présenté à la figure 5. (Foley et Guillemette, 2010.)

Figure 5
Processus de l'intelligence d'affaires.



Source: Foley et Guillemette (2010)

On observe cependant que la discipline de l'intelligence d'affaires ne présente pas d'emblée de code d'éthique pour soutenir ses professionnels en respectant des balises déontologiques et éthiques malgré les conseils de Schultz et al. (1994). En effet, ces derniers soulignent dans leur article la nécessité pour chaque entreprise d'avoir un code d'éthique, incluant la collecte et l'utilisation des données. Ils laissent ainsi entrevoir une problématique à considérer et régler. D'autres disciplines des affaires, telles que la comptabilité, le marketing ou encore les ressources humaines par l'entremise d'un ordre professionnel ou de regroupements officiels ont produits de tels cadres comme on peut le voir sur le site officiel de l'ordre des conseillers en ressources humaines agréés, ou bien sur celui des comptables professionnels agréés ou encore des professionnels en marketing. (CRHA, 2019 ; CPA, 2019 ; ACM, 2019)

L'absence de balises éthiques pour la gestion et l'analyse de données est mise en évidence dans le milieu académique et professionnel par Thomann et Wells (2013). Toutefois, les universités sont toujours à l'étape de réflexion sur les concepts et les enjeux de haut niveau plutôt que la définition des enjeux déontologiques et éthiques spécifiques à l'intelligence d'affaires et analytiques pour le praticien en dépit de l'augmentation du volume et de la variété des données accessibles pour le stockage et l'analyse dans les organisations.

À l'international, l'analyse des directives de deux grands regroupements d'intelligence compétitive relève l'importance de baliser l'usage des données en contexte d'affaires. Le SCIP en Amérique du Nord et de la FEPIE (Fédération des Professionnels de l'Intelligence économique remplacée en 2010 par le SYNFIÉ : Syndicat français de l'intelligence économique) en France offre des lignes de conduite régissant le comportement de leurs adhérents, principalement sur le processus de collecte des données externes à l'organisation (SYNFIÉ, 2018; SCIP, 2018). On constate que leur perspective fait cependant état d'une perspective différente de l'IA&A principalement axée sur le mode de gestion des interactions humaines.

On remarque aussi que malgré l'effort des différents acteurs du milieu de l'information, la question de la gestion et l'exploitation éthiques des données reste préoccupante. Face à cela, l'Union européenne a adopté le 25 mai 2018 le règlement

général sur la protection des données (RGPD) personnelles de ses ressortissants (Commission européenne, 2018), laquelle s'applique à l'ensemble des entreprises faisant affaire avec un citoyen ou résident d'un pays membre de l'Union européenne.

Par la formalisation de la protection des données personnelles des individus, cette loi permet d'entrevoir la nécessité de baliser l'utilisation des données personnelles par les spécialistes de la gestion et de l'analyse de données en entreprise en ciblant le processus même d'intégration et de gestion des données.

En effet, le règlement européen se concentre sur la collecte et le stockage des données personnelles et ne couvre pas directement ni l'exploitation ni l'utilisation des données au-delà d'un principe de divulgation de l'usage général des données par l'organisation (figure5: encadré rouge). Par l'intégration des données d'entreprise et le développement de nouveaux outils d'analyse, l'intelligence d'affaires et l'analytique permet l'émergence de nouvelles réalités, comme la déanonymisation et la réidentification. De ce fait, aujourd'hui le croisement d'une quantité suffisante de jeux de données anonymisés peut permettre sous certaines conditions l'identification, volontaire ou fortuite, de façon unique des personnes, des objets ou des phénomènes sous étude. De surcroît, le grain de popularité de l'intelligence artificielle en modélisation d'affaires, et particulièrement les modèles de type boîte noire où les mécanismes d'inférence sont implicites, et peut sous certaines conditions intégrer des variables jugées sexistes, racistes ou discriminatoires bien que significatif dans un modèle explicatif ou prédictif.

On constate qu'il n'existe pas de code d'éthique de référence pour encadrer spécifiquement l'exploitation et l'analyse des données massives, et ce malgré des pratiques répandues de collecte et de stockage de données dans les entreprises. Le cadre théorique proposé par Thomann et Wells (2013) établit trois activités sujettes à des considérations d'ordre éthique, notamment :

- a. La manière dont les données et les renseignements sont collectés ;
- b. La manière dont les données et les renseignements sont utilisés ;
- c. La manière dont on guide les conduites individuelles et organisationnelles à travers l'utilisation des données et des renseignements.

L'absence de balises formelles invite les spécialistes de l'intelligence d'affaires et de l'analytique à réfléchir sur la façon de combler l'absence de préceptes et de balises éthiques dans leur domaine. Le besoin est particulièrement criant dans les organisations pratiquant l'analyse de larges volumes de données détaillées provenant d'une multitude de sources sous des formats variés à des fins d'optimisation des pratiques d'affaires ou au développement d'avantages stratégiques.

Étant donné la définition du RGPD en Europe, la présente étude cible les deux dernières activités établies par Thomann et Wells (2013) dans leur cadre théorique. **En effet, l'étude s'intéresse à la perspective des praticiens sur les éléments prioritaires qui doivent donner naissance à la définition d'un encadrement de l'exploitation déontologique et éthique des données.**

2. Importance du sujet de recherche

L'IA&A est plutôt une discipline récente et nécessite encore de la recherche et du temps pour arriver à maturité. (Dussureaul, 2015.) À titre d'exemple, ce dernier cite le cas de l'intelligence compétitive qui a commencé à produire des analyses quantitatives et qualitatives seulement dans les années 1990. Ceci met en évidence son caractère plutôt jeune comparativement aux autres métiers du domaine d'affaires, dont la comptabilité, le marketing, la gestion des ressources humaines où leur maturité peut se voir notamment au travers de l'organisation de la profession autour d'un regroupement, tels l'ordre des comptables agréés, l'association des professionnels du marketing ou encore l'ordre des professionnels des ressources humaines. Les professionnels de ces regroupements sont soumis à des codes de déontologie régissant leur fonctionnement au quotidien et le non-respect est sujet à des sanctions. Baliser leur corps de métier permet non seulement de valoriser leur travail, mais aussi de légitimer leurs actions et de protéger à la fois les clients et les professionnels dans un cadre plutôt gagnant-gagnant. À l'opposé, ne pas encadrer de manière éthique la pratique de l'intelligence d'affaires et de l'analytique tel que mis à l'évidence par Thomann et Wells (2013) ouvrirait vers une avenue désavantageuse pour les entreprises qui pourraient voir : a) leur image de marque se ternir, b) une rupture de confiance dans

leur relation avec les clients, c) un manque à gagner financier dans leur revenu et bien d'autres conséquences.

Toutefois, il est difficile, voire impossible, d'établir aujourd'hui, sur la base de la littérature existante un code d'éthique ou de déontologie qui régit le métier de l'intelligence d'affaires et analytique et qui donne une ligne de conduite aux professionnels de ce domaine. Ceci révèle à nouveau la pauvreté des réponses académiques sur la question. En effet, une recherche dans les banques de données scientifiques de l'Université de Sherbrooke démontre l'absence de documentation sur le sujet¹. À titre d'exemple,

- Une recherche des mots clés « Ethics » et « Business Intelligence » sur ABIInform retourne un total de 8 articles écrits entre 1994 et 2019. Par comparaison, le nombre de résultats passe à 35 lorsque les critères de recherche sont modifiés pour prendre en compte le concept voisin d'intelligence compétitive avec « Ethics » et « Competitive Economic ».
- Dans la banque de données IGI, un total de 15 livres et de chapitres traitent de la question d'éthique en intelligence compétitive sont retrouvés avec le mot clé « Ethics » and « Competitive Economic », mais rien en intelligence d'affaires directement.
- Enfin aucun article n'apparaît dans la banque de données JStor avec les mots clés « Ethics » et « Business Intelligence ».

Le constat est étonnant, voire alarmant. Ceci pose l'hypothèse que la question de la déontologie et de l'éthique dans le domaine d'intelligence d'affaires et l'analytique n'est pas encore une préoccupation centrale des académiciens et des professionnels, et ce malgré l'engouement pour l'exploitation des données massives en entreprise.

Afin de permettre une bonne compréhension de la nécessité de cette étude, un parallèle est fait entre la pratique de l'intelligence d'affaires et l'analytique et la

¹ Notons que les chiffres indiqués ici sont les résultats de la recherche en Juillet 2018 au moment où nous définissions le cadre conceptuel de la présente étude.

conduite de véhicule. Imaginez quel chaos serait la conduite automobile en l'absence du code de la route : chacun déciderait impunément de faire ce qui lui convient et le nombre d'accidents de la route serait innombrable, les cas de décès suite aux accidents seraient inimaginables et les coûts financiers relatifs aux dégâts ne seraient pas des moindres. Il en est de même dans le domaine de l'intelligence d'affaires et l'analytique; en l'absence d'un code d'éthique et de déontologie régissant la pratique du métier, on se dirige irrémédiablement vers le chaos de nature différente à celle relative à la conduite routière certes, mais dont les dégâts sont coûteux et néfastes ; d'où la nécessité d'agir pour mettre sur pieds les fondements qui mèneront vers une pratique saine et salubre pour tous les acteurs de l'intelligence d'affaires et l'analytique.

Aussi, les scandales autour de l'exploitation des données comme cité à l'introduction du présent document contribuent à souligner l'importance de se pencher sur l'encadrement éthique de la pratique de la discipline d'intelligence d'affaires et analytique, particulièrement en ce qui a trait à l'exploitation des données par des professionnels du domaine et justifient une fois de plus la présente recherche tout en mettant l'accent sur le fait que le sujet traité par la présente étude a un impact réel tant pour les professionnels que pour les entreprises et leurs clients.

Ainsi est démontrée l'importance de cette recherche. La section suivante traite la problématique de recherche et donne les raisons pour lesquelles la présente étude a été conduite.

3. Problématique de recherche

L'entreprise utilise de plus en plus les données auxquelles elle a accès pour l'analyse de ses activités d'affaires afin d'en tirer un avantage concurrentiel. Au moment où la protection des données personnelles représente un défi majeur pour les gouvernements et les entreprises, l'éthique devient une préoccupation centrale. En contexte de responsabilité sociétale des entreprises (RSE), une mauvaise gestion ou utilisation des données peut devenir une occasion de briser des relations d'affaires entre compagnies notamment entre les investisseurs et les PME, entre les

entreprises et leurs clients ou encore entre employeur et employé. Par exemple, lorsque les méthodes qui ont contribué à la possession et à l'exploitation des données jetteraient un quelconque opprobre sur l'image de marque d'une des parties mis en cause.

Cette situation représente une véritable problématique dans la pratique de l'intelligence d'affaires et l'analytique et dans les relations d'affaires entre parties prenantes. Le fait que la profession de l'intelligence d'affaires et de l'analytique ne soit pas balisée par un code d'éthique et de déontologie contrairement aux autres métiers du domaine des affaires constitue la raison principale pour laquelle cette recherche est conduite.

Face à cela, la présente étude pose comme problématique **la détermination des priorités à considérer lors de l'élaboration d'un code d'éthique et de déontologie dans le domaine de l'IA&A** afin de prémunir l'organisation face aux problèmes que pourrait engendrer une pratique non éthique de l'intelligence d'affaires et de l'analytique.

Notons que l'intégration du terme éthique dans le domaine des affaires et particulièrement en intelligence d'affaires et analytique est très récente, comme le témoigne le rapport de décembre 2018 de Gartner (2018) où il est mis en évidence que l'éthique en intelligence d'affaires et analytique fait partie des tops 10 tendances managériales de 2019 (Gartner, 2018). Aussi, la discipline de l'intelligence d'affaires et analytique en elle-même est une discipline récente, comme en témoigne le nombre très limité des articles existant dans la littérature d'une part et d'autre part le fait que c'est seulement en 2010 que Foley et Guillemette (2010) en donnaient une définition précise. On constate par ailleurs qu'il est difficile de définir de manière universelle ce qui est éthique en intelligence d'affaires aujourd'hui et ce qui ne l'est pas au travers de la littérature. Ce manque constitue une autre raison pour laquelle cette étude est conduite. Aussi, le fait que l'évolution technologique a atteint un niveau où presque toutes les compagnies petites et grandes se trouvent à collecter et manipuler au quotidien un volume important de données, parfois sans avoir la qualification nécessaire pour un usage sécuritaire et éthique de ces données,

est inquiétant et nécessite une action d'où une autre raison d'être de la présente recherche.

Dans cette section, la problématique de cette recherche a été posée, ainsi que la motivation de la conduite de celle-ci. La section suivante traite l'objectif de la recherche tout en faisant ressortir les sous-objectifs de recherche.

4. Objectifs de recherche et résultats anticipés

Puisque l'IA&A est le domaine responsable de l'intégration des données des autres domaines d'affaires de l'entreprise, il présente des spécificités qui requièrent une attention et un traitement dédiés. Il convient de considérer l'encadrement de cette profession de manière conséquente et d'établir des balises spécifiques à la réalité des spécialistes de la gestion et de l'analyse des données.

L'objectif de cette étude est donc

- De poser les bases conceptuelles pour l'élaboration d'un code d'éthique et de déontologie adapté, répondant aux impératifs des professionnels du domaine en l'absence de balises actuelles encadrant l'exploitation éthique des données malgré l'engouement de plus en plus croissant pour les données.
 - Déterminer les principes fondateurs selon la perspective des professionnels
 - Déterminer le mode de gestion à privilégier par ces mêmes professionnels.

Pour atteindre cet objectif, la question de recherche de cette étude est formulée ainsi qu'il suit : **quels sont les enjeux que doit aborder un code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique selon les priorités perçues par les professionnels et les gestionnaires du domaine ?** En effet, à travers cette question, il est question d'établir les priorités à considérer lors de l'élaboration d'un code d'éthique dans le domaine ceci afin de se prémunir face aux problèmes que pourrait engendrer une pratique non éthique de l'intelligence d'affaires et de

l'analytique. Cette initiative aboutira également à des conclusions concrètes que pourront utiliser les entreprises de toute taille (petite, moyenne ou grande) pour définir leur code d'éthique et de déontologie autour de la gestion et l'exploitation des données pour leurs employés. À terme, le respect d'un code d'éthique et de déontologie en intelligence d'affaires et analytique offrira des bénéfices concrets à l'entreprise, notamment la préservation de l'image de marque de la compagnie, la confiance renforcée dans la relation avec la clientèle, le respect des lois, la non-discrimination dans le traitement des profils clients ou employés.

Dans cette section problématique, on a défini la problématique à travers les éléments comme la mise en contexte et la pertinence de cette recherche. Ici on a rappelé le fait que d'autres disciplines du domaine d'affaires sont encadrées par un code d'éthique et de déontologie, mais que ce n'est pas encore le cas pour l'intelligence d'affaires et l'analytique. On a mis en évidence la nécessité de l'encadrement de la discipline étant donné l'engouement de plus en plus croissant observé au niveau des organisations pour l'intelligence d'affaires et l'analytique et les enjeux qu'une pratique non encadrée représente. On a également présenté l'objectif de cette recherche et énoncé la question de recherche. Pour atteindre les objectifs de cette problématique, une méthodologie de travail est adoptée et permet d'avoir une démarche scientifique. La section suivante va présenter la méthodologie choisie pour cette étude, la mise en application ainsi que ses avantages et ses inconvénients.

CHAPITRE 4 : MÉTHODOLOGIE

Cette section présente la méthode utilisée pour cette recherche. Dans un premier temps, on décrit la méthode DELPHI en contexte d'application pour cette étude. Ensuite, on détaille la démarche du choix des répondants. Enfin, on fait état des limites de la méthode.

1. Choix de la méthode

Cette étude est de nature exploratoire en raison du manque de connaissances scientifiques sur le domaine de l'éthique en intelligence d'affaires et analytique. Plusieurs techniques permettent de conduire ce type d'étude (Ingham, 2016). La méthode DELPHI, privilégiée dans cette recherche, repose sur l'idée que les professionnels de l'intelligence d'affaires et de l'analytique, qui sont au premier rang des données, qui manipulent au quotidien les données et qui sont exposés aux enjeux éthiques entourant leurs activités, ont une position privilégiée pour se prononcer sur l'éthique de leur travail. Leur compréhension intime de l'exploitation des données leur confère une perspective inédite et éclairée.

La méthode DELPHI encore appelée méthode de Delphes est une méthode qui vise à organiser une consultation d'experts sur un sujet précis (Booto, Bernard et Plaisent, 2011). La méthode DELPHI a pour but de mettre en évidence des convergences d'opinions et de dégager certains consensus sur des sujets précis, grâce à l'interrogation d'experts, à l'aide de questionnaires successifs (Baillette, Fallery et Girard, 2013). Elle a été choisie pour cette étude, car elle permet de sonder les experts d'un domaine précis à un moment où peu d'études sont faites sur le sujet. Elle constitue un instrument d'aide à la décision pour explorer un sujet d'étude, notamment en phase exploratoire de recherche, pour construire des modèles ou encore pour réaliser les prévisions (Baillette, Fallery et Girard, 2013). Le choix de la méthode DELPHI est donc justifié par le fait qu'elle aide à consigner les avis des experts dans un domaine précis (Booto, Bernard et Plaisent, 2011) ; elle permet de recueillir les opinions à distance, sans que les experts soient tenus de se rencontrer (Booto, Bernard et Plaisent, 2011)

donnant ainsi la possibilité aux répondants de s'exprimer librement en évitant toute influence basée sur le statut ou la personnalité des répondants (Baillette, Fallery et Girard, 2013). Le dialogue et les échanges entre experts se font par correspondance interposée. Un analyste (facilitateur) joue le rôle d'agent centralisateur et s'occupe du bon déroulement de la méthode. La méthode DELPHI limite le nombre de participants à l'étude entre 7 et 18 experts (Bootto, Bernard et Plaisent, 2011). Enfin, l'analyse d'opinions des experts anonymes préalablement identifiés confère à la méthode DELPHI un avantage sur d'autres méthodes de prise de décision en groupe (Rohrbaugh, 1979, cité par Okoli et Pawlowski, 2004).

2. Description de la méthode DELPHI

La méthode d'enquête DELPHI, conçue en 1960 par Olaf Helmer à la Rand Corporation (Okoli et Pawlowski, 2004), vise à organiser la consultation d'experts sur un sujet précis (Dalkey et Helmer, 1963). Ici, le terme « expert » renvoie aux personnes ayant une bonne connaissance pratique, politique, légale ou administrative d'un sujet précis avec une légitimité suffisante pour exprimer un avis représentatif du groupe d'acteurs auquel ils appartiennent. Par conséquent, la méthode n'est pas réservée seulement aux autorités scientifiques, mais peut s'utiliser dans n'importe quelle étude (Landeta, 2006).

Elle consiste en un processus itératif qui permet une présentation anonyme d'un ensemble de jugements émanant de personnes informées. Elle a la particularité de poser une série de questions à un cercle permanent d'experts plusieurs fois de suite (Rowe et Wright, 1999).

En 1975, Linstone et Turoff (1975) ont proposé une définition qui selon eux résume la méthode ainsi que ses objectifs. Pour ces auteurs, « La méthode de Delphes peut être caractérisée comme une méthode pour structurer un procédé de communication de groupe de sorte que le processus soit efficace en permettant à un groupe d'individus, dans l'ensemble, de traiter un problème complexe ».

La mise en œuvre de la méthode DELPHI traditionnelle passe par les étapes suivantes selon Baillette, Fallery et Girard (2013) :

- Définir avec rigueur et précision l'objet sur lequel portera la méthode. L'objet correspond au problème à examiner par les experts et les questionnements liés à ce problème. Il est important à cette étape de bien définir l'objet afin d'éviter que les experts s'engagent dans un processus dont le thème évolue au fur et à mesure. Aussi une bonne définition de l'objet permettra de bien préparer la rédaction du questionnaire et de bien choisir les experts.
- Choisir les experts. Il s'agit d'un élément essentiel qui conditionne la qualité des résultats (Adler et Ziglio, 1996 ; Bolger et Wrigth, 1994). Puisqu'il ne s'agit pas d'une simple enquête d'opinion, la validité d'une enquête DELPHI ne dépend pas de l'échantillonnage, mais du savoir, de la compétence et surtout de la coopération intentionnelle des experts sélectionnés (Baillette, Fallery et Girard, 2013). Il s'agit donc de sélectionner les experts en fonction de leur expérience, de leur familiarité avec l'objet de l'étude, et de la connaissance des caractéristiques de l'objet (Baillette, Fallery et Girard, 2013). D'après Booto, Bernard et Plaisent (2011), la procédure de sélection comprend quatre étapes, soit l'élaboration des critères de sélection d'experts, l'élaboration de la liste des d'experts potentiels et l'attribution d'un numéro pour assurer l'anonymat, le contact avec les experts sélectionnés et finalement l'invitation des experts à participer à l'enquête en utilisant le courriel ou le fax. Afin d'éviter une remise en question future du processus DELPHI, il est important de veiller à :
 - Une bonne représentativité des experts choisis par rapport l'objet visé ;
 - Leur excellente connaissance de l'objet sur lequel porte le DELPHI ;
 - Leur indépendance par rapport à des pressions commerciales, politiques ou autres, afin d'assurer que les avis ne soient point influencés ni biaisés.
 - Leur légitimité par rapport au groupe d'acteurs qu'ils pourraient représenter ;
 - Leur disponibilité durant le processus DELPHI
- Élaborer un questionnaire. Les questions doivent être ciblées, précises et éventuellement quantifiables.
- Administrer le questionnaire et traiter les réponses. Le premier questionnaire, qui sert de base est envoyé individuellement aux experts par courrier (physique ou électronique) afin d'éviter les phénomènes d'influence liés au groupe. Ce

questionnaire sera par la suite enrichi, à chaque tour, par des résultats et commentaires générés précédemment.

Au deuxième tour de questionnaire, les experts reçoivent les résultats du premier tour et doivent à nouveau se prononcer sur le questionnaire, en ayant maintenant l'opinion du groupe consulté. Si leur nouvelle réponse dévie fortement de la moyenne du groupe, ils doivent la justifier. Au troisième tour, on informera les experts des résultats du deuxième tour ainsi que des commentaires justifiant les opinions déviantes. Les experts seront à nouveau invités à répondre au questionnaire, mais aussi à commenter les opinions déviantes. Le quatrième et dernier tour livrera aux experts toute l'information récoltée au cours des tours précédents et leur demandera de répondre une nouvelle fois au questionnaire. Ce quatrième tour donnera les réponses définitives : les opinions consensuelles médianes et la dispersion des opinions autour de cette médiane, cette dispersion pouvant être interprétée à l'aide des justifications et commentaires recueillis auprès des experts. La figure 6 récapitule les différentes étapes précédentes, nécessaires à l'exécution théorique et idéale d'une enquête DELPHI à savoir le processus de sélection des experts et le processus d'administration du questionnaire.

Cependant, il peut arriver qu'on ne se rende pas au quatrième tour comme l'indique le tableau 1 suivant où Baille, Fallery et Girard (2013) ont recensé quelques études menées à l'aide de la méthode de Delphes et où le consensus a été établi dès le deuxième tour ou le troisième tour de l'enquête.

Tableau 1

Études réalisées à l'aide de la méthode de Delphes.

IS/IT Study	Delphi Focus	Rounds	Sample Size
Niederman, Brancheau, & Wetherbe, (1991)	Survey senior IS executives to determine the most critical IS issues for the 1990s.	3	114, 126 & 104
Duncan (1995)	Identify and rank the critical elements of IS infrastructure flexibility.	2	21
Brancheau, Janz, & Wetherbe (1996)	Survey SIM members to determine the most critical IS issues for the near future.	3	78, 87 & 76
Nambisan et al. (1999)	Develop a taxonomy of knowledge creation mechanisms.	3	11
Scott (2000)	Rank technology management issues in new product development projects	3	20
Wynekoop & Walz (2000)	Rank the most important characteristics of high performing IT personnel.	3	9
R. Schmidt, Lyytinen, Keil, & Cule (2001)	Identify and rank software development project risks: an international comparative study.	3	Finland: 13, 13, & 13 Hong Kong: 11, 11 & 9 USA 21, 21 & 9
Keil, Tiwana, & Bush (2002)	Rank software development project risks.	3	15, 15 & 10
Brungs & Jamieson (2005)	Identify and rank computer forensics legal issues.	3	11

Source : (Baillette, Fallery et Girard, 2013)

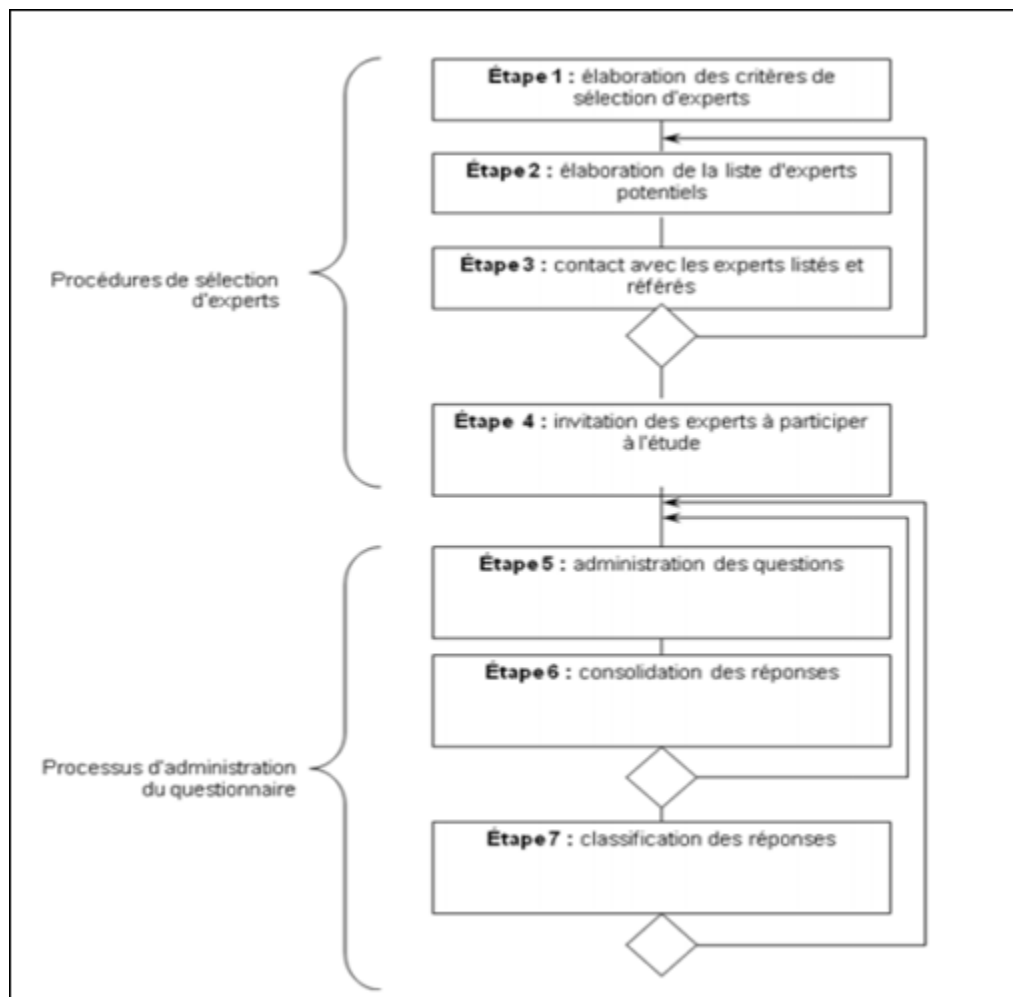
Tableau 2

Études plus récentes réalisées à l'aide de la méthode de Delphes.

Auteurs	But	Tours	Panel
(Gonzalez et al., 2006)	Explorer comment les managers SI évaluent les problématiques clés d'externalisation en Espagne.	2	1 ^{er} tour : 4416 managers SI de firmes espagnoles >357 réponses ; 2 ^{ème} tour : 49 > 15 réponses.
(Yeoh et al., 2008)	Identifier les facteurs clés de succès concernant le management de la mise en œuvre de systèmes de Business Intelligence.	3	15 BI system experts.
(Singh et al., 2009)	Identifier et surmonter les défis de la mise en œuvre d'un bureau de gestion de projet.	3	22 managers de projets TI expérimentés.
(McFadzean et al., 2011)	Identifier les processus associés au renforcement de l'alignement entre la sûreté de l'information, les SI et les stratégies d'entreprise.	3	36 experts de la sécurité des SI : praticiens et académiques.
(Bootoo Ekiona et al., 2011)	Identifier les concepts clés des capacités organisationnelles spécifiques de la gestion des connaissances	3	22 experts > 16 réponses.
(Kasiri et al., 2012)	Développer un « balanced scorecard » pour l'étiquetage RFID dans le secteur de la distribution.	2	10 consultants et managers seniors.

Source : (Baillette, Fallery et Girard, 2013)

Figure 6
Principales étapes de la méthode DELPHI



Source : (Booto, Bernard et Plaisent, 2011).

3. Avantages de la méthode

La description de la méthode DELPHI, telle que faite dans les paragraphes précédents, laisse entrevoir plusieurs avantages notamment :

- Elle aide à consigner les opinions d'experts dans un domaine précis;
- Elle permet de recueillir les opinions à distance, via Internet ou par la télécopie, sans que les experts soient tenus de se rencontrer;

- Elle limite le nombre de participants à l'étude entre 7 et 18 experts (Paliwoda, 1983, cité par Okoli & Pawlowski, 2004, puis repris par Booto, Bernard et Plaisent, 2011), facilitant du même coup le travail d'identification et de sélection d'experts;
- Elle est flexible dans sa conception et dans son administration du questionnaire;
- Elle aide à obtenir avec certitude un consensus à l'issue des questionnaires successifs;
- Elle facilite la rétroaction contrôlée, qui consiste en une série d'étapes au cours desquelles un sommaire de l'étape précédente est communiqué aux participants, permettant à ces derniers, s'ils le désirent, de réviser leurs jugements antérieurs.
- Finalement, l'analyse d'opinions d'experts anonymes préalablement identifiés confère à la méthode DELPHI l'avantage sur d'autres méthodes de prise de décisions en groupe (Rohrbaugh, 1979, cité par Okoli & Pawlowski, 2004, puis par Booto, Bernard et Plaisent, 2011).

4. Limites de la méthode

Comme la plupart des techniques qualitatives utilisées en recherche exploratoire, la méthode de Delphes présente quelques inconvénients dont il importe de souligner ici. En effet, le taux d'abandon élevé reste une des limites primordiales (Baillette, Fallery et Girard, 2013), sans doute dû au nombre d'itérations, constituant un défaut majeur de cette méthode. La méthode est longue, coûteuse, intuitive plutôt que rationnelle et fastidieuse surtout dans son application traditionnelle tant pour l'analyste que pour les experts à cause des quatre itérations requises (Booto, Bernard et Plaisent, 2011). Aussi, seuls les experts qui sortent de la norme sont amenés à justifier leur position. Toujours d'après Booto, Bernard et Plaisent (2011), on peut considérer que l'opinion des déviants est en termes prospectifs plus intéressante que celle de ceux qui rentrent dans le rang. Enfin, une enquête DELPHI ne repose pas sur un échantillon statistique représentatif de l'ensemble de la population. Il s'agit plutôt d'un mécanisme de prise de décision de groupe qui requiert la participation d'experts qualifiés qui ont une compréhension claire de l'objet à l'étude.

5. Conception de l'enquête pour l'étude

Le déroulement de la présente étude DELPHI a respecté les sept étapes évoquées à la section précédente. Ainsi, après avoir élaboré le questionnaire, on a choisi un nombre d'experts répondant aux critères précédemment nommés. Ensuite, le questionnaire leur a été distribué une première fois. Puis, après une première analyse des réponses fournies, une autre itération du questionnaire a été administrée en demandant cette fois-ci aux experts de revoir leurs réponses (opinions) originales ou de répondre à certaines questions précises en fonction de la rétroaction (*feedback*) venue des autres participants à l'étude. Au terme de cette seconde itération, le consensus était observé en majorité, ce qui a permis d'arrêter l'enquête. Au premier tour, les concepts énoncés par la littérature comme des concepts liés à l'utilisation éthique des données en intelligence d'affaires et analytique ont été présentés aux experts en guise de première question. Il leur a été demandé de dire si d'après leur expérience, ces concepts sont effectivement liés à l'utilisation éthique des données en intelligence d'affaires et analytique, puis pour chacun des concepts, il leur a été demandé de dire quel sens ils accordaient à chaque concept. En guise de deuxième question, il leur a été demandé de proposer d'autres concepts qui, selon eux, sont liés à l'utilisation éthique des données en intelligence d'affaires et analytique puis de donner un sens à chaque concept proposé. En guise de troisième question, il leur a été demandé de classer les concepts de la question 1 et 2 d'après leur expertise selon un ordre de priorité allant du plus prioritaire au moins prioritaire. En guise de quatrième question, les experts devaient selon leur expertise, dire qui dans l'entreprise doit veiller à l'effectivité d'un tel code. Enfin en guise cinquième question, les experts devaient proposer un mécanisme de contrôle de l'effectivité d'un tel code. Au deuxième tour, un rapport du premier tour du questionnaire DELPHI a été fourni aux experts participants. Ainsi, conformément à ce rapport, il a été demandé aux experts de se prononcer sur des questions spécifiques autour des points de consensus et de divergences d'opinions.

En somme, les étapes suivies par cette recherche ont été : 1) la définition des critères de sélection d'experts-participants; 2) l'élaboration de la liste d'experts; 3) les contacts avec les experts listés et référés; 4) l'invitation des experts à participer à l'étude; 5) l'administration du questionnaire; 6) le dépouillement complet et la consolidation des réponses selon différents angles de vue et présentation des résultats; et

7) l'élaboration de synthèses et la classification des caractéristiques des concepts clés. Ainsi, cette enquête DELPHI comporte deux parties : les étapes relatives à la procédure de sélection d'experts et le processus d'administration du questionnaire précédé par l'obtention du certificat d'éthique (Annexe 13).

a) La Procédure de sélection des experts

Il est question ici de reprendre chacune des étapes nécessaires à la sélection des experts afin de rendre compte de ces divers éléments dans la présente étude.

Étape 1 : Critères de sélection d'experts

Lors de cette étape, deux groupes ont été ciblés. Les principaux critères retenus pour la sélection des experts sont que ces derniers doivent être essentiellement dotés de connaissances théoriques et pratiques dans le domaine de l'intelligence d'affaires et l'analytique et posséder au minimum trois années d'expériences dans leur domaine d'expertise. De plus, ils doivent encore être en fonction à leur poste au moment de la participation à l'enquête. Ainsi, cette enquête s'adresse à des professionnels en exploitation des données, incluant des analystes de données dans les lignes d'affaires, les scientifiques des données, les spécialistes en analytique d'affaires, les statisticiens ou toute autre personne occupant des fonctions similaires d'une part et à des gestionnaires d'équipes d'exploitation des données ou de gestion des données en intelligence d'affaires et analytique d'autre part. Le choix des professionnels en exploitation des données est justifié par le fait que ces derniers interagissent au quotidien avec la donnée et sont souvent confrontés aux défis d'ordre éthique. Quant au choix des gestionnaires d'équipe, il est justifié par le fait que les gestionnaires ne sont pas au contact direct avec les données, mais leur responsabilité est engagée du fait même de leur rôle. Enfin le choix de deux populations encore en fonction dans leur poste respectif est justifié par le fait de pouvoir comparer les résultats de l'enquête en fonction des rôles et responsabilités vis-à-vis de la donnée ; de plus, en étant encore en poste, ils possèdent des connaissances récentes et à jour et peuvent de ce fait fournir une réponse claire et factuelle. Ainsi, l'étude s'intéresse à des corps d'emploi présentant un objet d'étude commun, mais ayant des rôles et des responsabilités différentes vis-à-vis de la donnée.

Les candidats seront d'abord invités à participer à l'étude par l'entremise de regroupements professionnels établis dans le domaine ceci permet de réduire le risque. En l'absence d'un ordre professionnel ou d'une organisation officielle chapeautant la pratique en intelligence d'affaires les groupes suivants ont été privilégiés :

- Réseau BI du Réseau Action TI (<https://www.actionti.com/>), le directeur du Réseau BI est présentement chargé de cours en intelligence d'affaires pour l'Université de Sherbrooke.
- Réseau « Data Driven Montréal », réseau de professionnels du domaine dirigé par un collaborateur en intelligence d'affaires et l'analytique à l'Université de Sherbrooke.
- Groupe LinkedIn « Intelligence d'affaires région de Québec ».

Étape 2 : Élaboration de la liste d'experts

Après la consultation des groupes mentionnés à l'étape 1, une liste de 167 (soit 107 professionnels en exploitation des données et 60 gestionnaires en IA&A) experts a été dressée en format Excel contenant le numéro confidentiel attribué à l'expert pour garantir son anonymat, son adresse électronique, son numéro de téléphone, le lien vers son profil LinkedIn, le nom de la compagnie actuelle pour laquelle il travaille, le poste actuel (titre) ainsi le nombre d'années d'expérience dans cette compagnie. Le nom de la compagnie précédente, le poste dans la compagnie précédente ainsi que le nombre d'années d'expérience dans la compagnie précédente. Cependant, avant d'élaborer la liste des répondants, une demande d'approbation de l'étude a été faite le 5 septembre 2018 auprès du comité éthique de l'université, car l'étude porte sur les êtres humains. Le traitement a duré 14 jours et a abouti à une approbation conditionnelle le 19 septembre 2018. La correction de la demande initiale a été faite conformément aux exigences mentionnées dans la lettre d'approbation conditionnelle puis soumise de nouveau le 21 octobre 2018. Le traitement de cette nouvelle demande a duré 1 mois et 2 jours et l'approbation finale (Annexe 13) a été rendue le 23 novembre 2018. À partir de cette date, on pouvait constituer la liste et contacter les potentiels répondants.

Étape 3 : Contact avec les experts listés et référés

Au cours de cette troisième étape, les experts listés à l'étape 2 ont d'abord été contactés par courriel (voir un exemple de lettre adressée à un expert à l'Annexe 1). Pour ceux dont nous n'avions pas d'adresse courriel, un message de sollicitation à leur a été envoyé sur LinkedIn en sollicitant leur intérêt pour l'étude (voir l'exemple de message de sollicitation adressé à l'expert en annexe 3). Lorsque ce dernier prenait connaissance du message et fournissait son courriel, la liste Excel était alors mise à jour. Comme le recommande la méthode DELPHI, il a été demandé à certains experts participants de recommander à d'autres experts répondant aux critères de l'enquête. Sur un total de 167 experts potentiels visés, 99 d'entre eux (soit 61 Professionnels et 38 gestionnaires), tous établis au Canada dans la région de Montréal et de Québec ont accepté de participer à l'enquête DELPHI, ce qui dépasse largement le minimum de 7 participants suggéré habituellement par une telle démarche méthodologique.

Étape 4 : Invitation des experts à participer à l'étude

La quatrième étape a consisté à inviter chacun des 99 experts à prendre part à l'enquête et à leur expliquer les objectifs de l'étude, les procédures à suivre et les devoirs liés à leur participation afin de garantir les chances de succès de l'étude. L'adresse courriel a été adoptée pour la réception ou l'envoi des questionnaires et des réponses. La lettre d'invitation était accompagnée d'une lettre de consentement (annexe 4) à participer où chaque participant était appelé à signer et le retourner avec sa réponse au premier questionnaire.

b) Processus d'administration du questionnaire

Au cours des étapes 5, 6 et 7 (Figure 6), il a été demandé aux experts de se prononcer sur les objets suivants : 1) les concepts liés à l'exploitation éthique des données ; 2) d'autres concepts à considérer 3) le classement de ces éléments ; 4) le responsable de l'effectivité d'une exploitation éthique des données ; 5) les mécanismes de contrôle. Deux tours de l'enquête DELPHI ont été nécessaires pour dégager un consensus autour des concepts liés à l'exploitation éthique des données, autour du responsable chargé de la garantie d'une exploitation éthique et pérenne des données,

autour du mécanisme et des outils de contrôle de l'effectivité d'une exploitation éthique des données.

Premier tour (Round 1) : envoi du questionnaire et consolidation des réponses des experts

Comme mentionné à la section précédente, le premier tour du questionnaire, qui s'est déroulé du 15 décembre 2018 au 23 février 2019, consistait à définir les concepts liés à l'utilisation éthique des données après avoir indiqué si selon leur expertise, ces concepts sont effectivement liés à l'utilisation éthique des données et devraient être pris en compte dans l'élaboration d'un code d'éthique et de déontologie destiné à l'exploitation des données. Au premier tour, les questions ont été élaborées et présentées aux experts en tenant compte des activités sujettes à des considérations éthiques telles que suggérées par Thomann et Wells (2013) à savoir *la manière dont les données et les renseignements sont collectés ; la manière dont les données et les renseignements sont utilisés et la manière dont on guide les conduites individuelles et organisationnelles à travers l'utilisation des données et des renseignements*. Par la suite, quatre questions supplémentaires ont été ajoutées tel qu'indiqué au début de cette section. Chaque expert devait répondre par oui ou par non à la première partie de la question 1, avec possibilité de donner des commentaires et pour les autres questions, une réponse ouverte avec également la possibilité de donner des commentaires était nécessaire.

Il faut dire qu'au lieu de deux semaines comme initialement prévu, le premier tour de l'enquête DELPHI a duré huit semaines et huit jours parce que la majorité des experts n'avaient pas répondu à temps et qu'il fallait atteindre un nombre suffisant de participants afin de s'assurer de la validité de l'enquête. Après ces huit semaines et huit jours, 22 (soit 11 par groupe de participants) des 99 experts avaient retourné leur questionnaire rempli. Pour mieux cerner les résultats du premier tour de l'enquête DELPHI, un rapport détaillé a été élaboré, rapport inspiré de la structure du questionnaire initial afin de faciliter la compréhension des experts lors du deuxième tour. Les résultats du premier tour ont fait ressortir les divergences de points de vue dans la compréhension des experts des concepts liés à l'utilisation éthique des données. Cinq questions ont été posées aux experts afin de recueillir leur opinion.

Deuxième tour (Round 2) : dépouillement, classification et consolidation des concepts liés à l'utilisation éthique des données.

Ayant constaté des divergences de points de vue dans la compréhension des experts sur six des sept concepts liés à l'utilisation éthique des données et sur certaines caractéristiques lors du premier tour de l'enquête, un deuxième tour a aidé à faire avancer le débat en vue d'obtenir un consensus autour de ceux-ci. Dans une enquête DELPHI, l'opinion négative ou divergente des experts est traitée avec beaucoup d'attention. Elle peut s'avérer plus intéressante que l'opinion positive, car elle permet d'identifier les points de divergence et de désaccord à la source d'une mauvaise compréhension des concepts et de problèmes éventuels à leur application. (Bootto, Bernard et Plaisent, 2011.) Pour le deuxième tour, on notera comme avantages que le phénomène peut se produire parce que la discussion avec les semblables augmente la certitude que quelques individus peuvent modérer leurs points de vue en fonction des points de vue opposés et que les influences argumentaires réciproques de certains membres sur les autres peuvent renforcer des orientations existantes. Le questionnaire du deuxième tour (annexe 8) a été distribué aux 22 experts encore répartis dans la région de Montréal et la région de Québec et qui ont répondu lors du premier tour. Une lettre d'accompagnement (annexe 7 rédigée par le directeur de cette étude) leur a aussi été envoyée par courriel pour les remercier de leur participation au premier tour, souligner l'importance de leur participation et les convier à participer au deuxième tour de l'enquête. Au lieu de deux semaines comme prévu au départ, le deuxième tour de l'enquête DELPHI s'est étendu sur 7 semaines (du 1er avril au 22 mai 2019), la majorité des experts n'ayant pas répondu à temps une fois de plus comme ce fût aussi le cas au premier tour.

Au terme du deuxième tour, 14 (soit 7 professionnels en exploitations des données et 7 gestionnaires) des 22 experts ont remis leur questionnaire rempli. Un autre rapport détaillé a été élaboré afin de mieux cerner les résultats de ce deuxième tour qui visait à atteindre ou à augmenter le consensus autour de certains concepts liés à l'exploitation éthique des données ainsi que leurs caractéristiques. Les résultats ont fait ressortir un consensus sur la définition et la compréhension de la majorité des concepts. Des questions ont de nouveau été posées aux experts afin de recueillir leurs opinions sur les concepts liés à l'exploitation éthique et les caractéristiques qui n'avaient pas obtenu

de consensus au premier tour. Les résultats du deuxième tour de l'enquête DELPHI ont aidé à consolider les réponses des experts en excluant les duplications et en unifiant la compréhension des concepts.

Un consensus jugé suffisant ayant été observé au terme de ce second tour. L'enquête a pris fin à ce stade et l'étape d'analyse détaillée des réponses a eu lieu.

La section suivante traite l'analyse détaillée des réponses suivant ces deux tours d'enquêtes.

Conclusion

Dans cette section, le choix de la méthode adoptée pour cette recherche, le choix des répondants, une description détaillée de la méthode ainsi que ses avantages et quelques limites de la méthode ont été présentés. Il a été également question de l'application de la méthode à la présente étude. Malgré les inconvénients connus, la méthode DELPHI convient pour cette étude et a permis l'atteinte des objectifs. Les résultats obtenus sont présentés dans la section suivante.

CHAPITRE 5 : RÉSULTATS ET DISCUSSION

Dans cette section, on présente premièrement les résultats du questionnaire DELPHI soumis aux professionnels de l'intelligence d'affaires et analytique. Deuxièmement, on effectue une analyse des résultats et on termine avec la présentation des limites de cette étude.

1. Description des résultats

Afin de déterminer les éléments prioritaires à considérer lors de l'élaboration d'un code d'éthique dédié à l'intelligence d'affaires dans un contexte d'analyse de données, une enquête DELPHI à deux tours a été soumise à deux groupes de professionnels du domaine. En effet, le premier tour d'enquête a débuté le 15 décembre 2018 et s'est clôturé le 23 février 2019. À la suite d'analyse préliminaire de ce premier tour d'enquête, un second tour a eu lieu du 1 avril 2019 (avec l'envoi des questionnaires) au 22 mai 2019 (avec le début de la compilation des différentes réponses). Ce second tour d'enquête était basé sur 4 constats découlant de l'analyse du premier tour. On va premièrement présenter l'analyse des résultats des professionnels en exploitation des données, ensuite l'analyse des résultats des gestionnaires d'équipe et enfin une analyse comparative des résultats des groupes clôtureront cette section.

a) Résultats des professionnels en exploitation des données

L'enquête pour cette étude s'est déroulée en deux tours. Comme indiqué dans la section précédente, au premier tour, les questions ont été élaborées et présentées aux experts en tenant compte des activités sujettes à des considérations éthiques telles que suggérées par Thomann et Wells (2013). Dans les lignes suivantes, on rappelle chaque question posée ainsi que la réponse des experts.

Question 1 du questionnaire du premier tour de l'enquête DELPHI portant sur les éléments liés à l'utilisation éthique des données.

« Une revue de la littérature scientifique et professionnelle a permis d'identifier des concepts liés à l'utilisation éthique des données en intelligence d'affaires et analytique. Les principaux concepts répertoriés incluent l'anonymat, la confidentialité, la sécurité, le secret, l'exactitude, l'intégrité et la transparence. Considérez-vous ces éléments comme déterminants à la définition d'un code d'éthique spécifique à l'intelligence d'affaires dans un contexte d'exploitation de données ?

Pour chaque terme, nous vous prions d'indiquer le sens que vous attachez au terme et une brève description de son importance dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives. »

Les réponses fournies par les experts à cette première question ont révélé un consensus à l'unanimité autour du concept « anonymat » en ce qui a trait au sens et à la description de ce concept comme **l'idée de ne pas dévoiler l'identité d'une personne en rapport avec un fait** (voir tableau 3). Cependant, on observe une forte divergence pour tous les autres concepts. Aussi, 6/11 répondants ne se sont pas prononcés sur la question de savoir si tous les concepts (éléments) sont déterminants à l'élaboration d'un code d'éthique et de déontologie spécifique à l'intelligence d'affaires et l'analytique dans un contexte d'exploitation de données. Parmi ceux qui se sont prononcés, on observe une divergence dans leurs réponses, 2/11 experts estiment que le concept est déterminant selon que la donnée exploitée est publique ou non et 3/11 experts sont totalement d'accord que tous les concepts énumérés sont déterminants dans l'élaboration d'un code d'éthique spécifique à l'intelligence d'affaires et l'analytique dans un contexte d'exploitation de données.

Tableau 3

Sens du terme Anonymat tels que compris par les scientifiques de données

Sens du terme anonymat tel que proposé par les experts		
Rendre impossible d'établir un lien en les données et la personne source qui les a fournies ou sur qui elles ont été recueillies	garantir l'impossibilité de faire le lien entre des données et les individus desquels elles proviennent.	certaines soient en mesure d'identifier des personnes internes
État d'une chose, d'une personne, dont on ignore le nom ou l'identité.	Ne pas être en mesure de lier la donnée à une personne	L'anonymat dans ce contexte concerne l'anonymisation des données sensibles.
Information qui est approvisionnée sans identifier exactement sa source, soit une personne, un groupe de personnes ou entreprise	la donnée est collectée sans aucune information pouvant permettre d'identifier la ou les personnes directement associées à cette donnée	l'impossibilité d'identifier les personnes physiques ou morales concernées par la donnée
Garder les données utilisables sans pouvoir identifier des informations personnelles en les remplaçant par des données de même nature. Par exemple, remplacer les noms et les prénoms à partir d'une banque de données aléatoire de noms et de prénoms.	Ne pas garder de données pouvant aider à l'identification d'un individu précis	

Source : compilation des réponses des scientifiques de données à la question 1

Question 2 du questionnaire du premier tour de l'enquête DELPHI portant sur les autres éléments déterminants dans l'élaboration d'un code d'éthique dédiée à l'intelligence d'affaires et l'analytique et qui ne faisaient pas partie de la liste proposée.

« Selon votre expertise, quels sont les autres concepts ou éléments qui doivent être pris en compte de manière prioritaire dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives ?

Pour chaque terme, nous vous prions d'indiquer le sens que vous attachez au terme et une brève description de son importance dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives. »

Les réponses du panel d'experts à cette seconde question concernant les autres concepts à considérer ont, dès le premier tour, révélé une divergence. Même si les concepts comme « consentement » « respect de la loi » reviennent à quelques reprises, on peut remarquer ici qu'il y a autant de concepts supplémentaires qu'il y a de répondants, chacun faisant une proposition selon sa compréhension de l'étude. On remarque cependant à ce niveau un fossé entre les éléments supplémentaires proposés et la liste des concepts trouvés dans la littérature. (Voir les éléments proposés par les experts en annexe 9.)

Question 3 du questionnaire du premier tour de l'enquête DELPHI portant sur le classement des concepts selon l'ordre de priorité du plus prioritaire au moins prioritaire.

« En vous basant sur votre expertise, placez les éléments retenus à la Question 1 et à la Question 2 par ordre de priorité du plus prioritaire au moins prioritaire pour l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives. »

Les réponses obtenues à cette troisième question ont révélé une grande divergence dans le classement des concepts. Ceci n'est pas surprenant si l'on s'en tient à la diversité de réponses observée aux questions 1 et 2. En effet, en se basant sur le classement fait par

les experts, on peut regrouper les concepts en trois grands groupes selon l'importance accordée aux concepts :

- Le premier groupe étant composé de **confidentialité, anonymat et sécurité.**
- Le deuxième groupe composé de **transparence.**
- Le troisième groupe composé de **secret, intégrité et exactitude.**

Le regroupement ne tient pas compte des éléments supplémentaires proposés par les experts étant donné que ces éléments étaient propres à chaque expert. Ces éléments hors sujet ont été portés en annexe 9 et n'ont pas été pris en compte dans la suite de l'étude. Il a été jugé pertinent de considérer uniquement les éléments pour lesquels tous les experts avaient donné effectivement un avis afin de rester cohérents dans les résultats de l'étude.

Question 4 du questionnaire du premier tour de l'enquête DELPHI portant sur la personne responsable de veiller à l'effectivité d'un tel code.

« À titre d'expert de l'exploitation de données, qui dans l'organisation devrait être responsable de veiller à l'effectivité, c.-à-d. à la mise en opération effective, mesurée et pérenne, de chacun des éléments retenus à la Question 1 et à la Question 2 ? »

Les réponses du panel d'experts à cette quatrième question du premier tour de l'enquête ont indiqué qu'il y a eu une divergence au niveau de la proposition des experts. Certains pensent qu'une équipe composée d'experts en intelligence d'affaires et des gens du métier devrait avoir cette responsabilité. D'autres proposent que cette responsabilité relève du chef officier de données, d'autres encore proposent des responsables différents pour chaque concept énuméré. On peut résumer les réponses en deux catégories : une catégorie centralisée où une personne (ou un comité) responsable serait désignée pour assurer cette responsabilité et une catégorie décentralisée où chacun serait responsable à un certain niveau de l'effectivité de ce code.

Une fois de plus, la divergence d'opinions observée à cette question est sans doute fortement liée à la compréhension des termes évoqués à la question 1 et proposée dans la littérature.

Question 5 du questionnaire du premier tour de l'enquête DELPHI portant sur le mécanisme de contrôle d'un tel code.

« À titre d'expert de l'exploitation de données, quels sont les mécanismes de contrôle de gestion à considérer pour chacun des éléments retenus à la Question 1 et à la Question 2 ? »

Les réponses obtenues à cette cinquième et dernière question ont révélé une certaine convergence dans l'approche au niveau des opinions des experts, mais une fois de plus une divergence au niveau des mécanismes. En effet, les experts proposent en majorité des mécanismes de contrôle pour chaque concept et on peut remarquer que ces propositions relèvent davantage de la gouvernance des données plutôt que du contrôle en soi (voir annexe 10). Cependant, 1 des 11 experts propose la création d'une commission nationale de norme éthique en matière d'analyse et d'exploitation des données et 2 des 11 répondants ont proposé un mécanisme plutôt général en deux approches statiques et dynamiques (tableau de bord permettant de surveiller les concepts considérés). À titre d'exemple, l'expert X déclare : *« Comme il y'a plusieurs concepts à surveiller avec un grand volume de données, d'accès et d'utilisation, je vois deux mécanismes : un qui est statique et l'autre est dynamique.*

Le mécanisme statique est l'audit, cette équipe doit s'assurer pour un projet/produit ou une application donnée que tous ces concepts sont respectés.

Le mécanisme dynamique est l'utilisation de la sécurité analytique qui inclut les techniques d'apprentissage automatique pour la détection en continu des anomalies ou le non-respect de ces concepts. »

Synthèse des résultats du premier tour de l'enquête DELPHI portant sur les éléments liés à l'exploitation éthique des données.

De l'ensemble des opinions récoltées chez les experts lors du premier tour de l'enquête DELPHI, ressortent les résultats suivants : 1) il y a un consensus autour du seul terme anonymat ; 2) il y a divergence d'opinions au sujet de la personne responsable de l'effectivité du code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique à la suite des commentaires formulés par les experts. Certains sont pour un contrôle interne à l'entreprise lorsqu'une minorité propose l'existence d'un

organisme étatique chargé de veiller à l'application du code au sein des entreprises ; 3) il règne une divergence au niveau du classement des concepts considérés ; 4) il y a absence de consensus autour du mécanisme de contrôle et, selon une minorité des experts, il serait indiqué de penser à une méthode de contrôle intégrant deux mécanismes dont l'un statistique relevant de l'audit et contrôle inopiné et l'autre dynamique, lié à l'apprentissage machine permettant de déterminer de façon continue et automatique les écarts dans le respect des éléments considérés.

Deuxième tour (Round 2) : dépouillement, classification et consolidation des concepts liés à l'utilisation éthique des données.

À la suite du premier tour de l'enquête, on a observé des divergences, ce qui n'est pas tout à fait surprenant du fait même de la nature de l'enquête. Ces divergences ont été formulées sous forme de constat et communiquées aux experts dans le but de souligner pour chaque question du premier tour, la divergence observée et ce qui est attendu au second tour.

Ainsi dans les lignes suivantes, les constats qui ont animé le second tour de l'enquête sont présentés. Ce second tour s'est déroulé sur une période de 7 semaines, soit du 1er avril au 22 mai 2019. ***Il faut noter qu'au deuxième tour de l'enquête, 7 experts ont retourné leur formulaire dûment rempli. Le nombre de participants passe donc de 11 à 7, justifiant ainsi pourquoi les résultats seront X/7 au lieu de X/11 comme précédemment ; où X représente le résultat constaté de l'élément analysé.***

Constat 1 découlant de la première question du premier tour de l'enquête : Disparité dans la compréhension des termes liés à l'éthique

Le premier constat porte sur la compréhension ou mieux le sens que les répondants donnent aux différents termes liés à l'éthique ; on a donc constaté une disparité dans la compréhension et on a soumis à tous les répondants les différents sens que les uns et les autres donnent aux termes et on veut les entendre de nouveau, suite à l'exposition des différentes compréhensions. On constate maintenant que les experts connaissent les avis les uns des autres, une convergence des opinions à l'issue de ce second tour d'enquête autour de la compréhension et du sens des termes soumis à leur attention. Cependant, certains experts ont encore des réserves notamment pour les

termes transparence, intégrité et secret. En effet, pour chacun des concepts, voici les réponses fournies par le panel :

- **Le concept confidentialité** : 7/7 des répondants s'accordent autour du sens « *la donnée est vue et utilisée uniquement par des personnes autorisées.* ».
- **Le concept sécurité** : 6/7 des répondants pensent que c'est « *l'action de protéger les données contre les accès non autorisés et contre toute altération de la donnée.* » Tandis que pour 1/7, il s'agit de « *mettre la donnée à l'abri de tout danger.* »
- **Le concept secret** : 5/7 des répondants pensent qu'il s'agit de : « *garder une discrétion dans le but de garantir la confidentialité* » ; 2/7 estiment que le concept secret est un synonyme au concept confidentialité.
- **Le concept exactitude** : 5/7 des répondants disent qu'il s'agit de la « *représentation fidèle de ce qu'elle est supposée effectivement représenter* » et 1/7 lui associe le sens de « *l'absence de biais ou d'erreurs* ».
- **Le concept intégrité** : 3/7 des répondants lui donnent le sens de « *donnée complète et fiable de bout en bout ainsi que l'intégrité des professionnels des données* » ; 1/7 limite son sens à « *la donnée complète et fiable de bout en bout, sans faire allusion à l'intégrité des professionnels.* » ; 2/7 le ramène à la fiabilité de la donnée tout au long de son parcours, sans toutefois tenir compte de la notion de complète et enfin 1/7 pense que ce concept est synonyme du concept exactitude.
- **Le concept transparence** : 4/7 lui donnent comme sens le fait « *d'expliquer la méthodologie et ne pas obstruer la vue d'ensemble des pairs lors du partage de la méthodologie et des résultats.* » 1/7 pense qu'il s'agit plutôt de « *divulguer tout ce qui est pertinent en lien avec la donnée; par exemple, le contexte d'utilisation, les transformations / altérations, la récolte, etc.* » ; 1/7 associe plutôt le sens « *d'informer la personne concernée par la donnée les fins d'usage* » enfin 1/7 pense qu'il s'agit de « *clarifier la manière dont les données sont collectées et utilisées* ».

Comme mentionné plus haut, on a un fort consensus sur la majorité des termes soit 5 des 7 termes soumis aux experts. À ce stade, on constate une grande évolution au niveau de la compréhension des répondants au second tour de l'enquête. Quelques-uns ont irrémédiablement changé le sens qu'ils donnaient à un ou plusieurs termes au premier tour de l'enquête ; sans doute influencé dans leur position par la compréhension des autres répondants. En effet, on observe qu'une grande majorité s'accorde dans la compréhension des termes confidentialité, sécurité, secret et exactitude. Les concepts intégrité et transparence cependant, présentent encore une disparité dans leur compréhension par l'ensemble des répondants. Il serait important de clarifier quels sens la littérature donne à ces concepts et davantage divulguer le sens que la littérature existante lui donne dans le but d'édifier les professionnels à ce niveau.

*Constat 2 découlant de la deuxième question du premier tour de l'enquête : **Éléments prioritaires dans la définition et la mise en œuvre d'un code d'éthique***

Le second constat lié à la classification des termes à la suite de la nouvelle compréhension faite par les répondants révèle que même à ce niveau, il y a eu un changement au niveau du classement. Le nouveau classement des éléments du plus prioritaire au moins prioritaire est le suivant :

<p>- Deuxième tour :</p> <ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité ○ Transparence ○ Exactitude ○ Intégrité ○ Anonymat ○ Secret 	<p>- Premier tour :</p> <ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité ○ Anonymat ○ Transparence ○ Secret ○ Exactitude ○ Intégrité
---	--

On peut vite remarquer qu'à l'exception des termes « sécurité » et « confidentialité » qui ont gardé leur position, tous les autres éléments ont bougé matérialisant l'évolution observée précédemment dans la compréhension des termes.

Ces classements ont été regroupés en trois catégories selon **l'importance accordée** à chaque terme par les experts comme suit :

Premier groupe d'éléments plus prioritaire regroupe	confidentialité (5/7 experts l'ont mis en 2 ou 3 ^e position), sécurité (5/7 experts l'ont mis en 1 ^{ere} position)
Deuxième groupe d'éléments plus prioritaire regroupe	transparence (3/7 experts l'ont mis en 1 ou 3 ^e position), Intégrité (3/7 experts l'ont mis en 1 ou 3 ^e position), et exactitude (3/7 experts l'ont mis en 2 ou 3 ^e position),
Troisième groupe d'éléments plus prioritaire regroupe	Secret (3/7 experts l'ont mis en 7 ^e position), Anonymat (4/7 experts l'ont mis en 6 ou 7 ^e position),

Constat 3 découlant de la troisième question du premier tour de l'enquête :
Gouvernance de l'éthique

Le troisième constat porte sur la gouvernance de l'éthique dans l'organisation. Il a été demandé aux répondants de se prononcer au sujet de la centralisation ou de la décentralisation de la gouvernance de l'éthique en intelligence d'affaires et analytique. *Il faut signaler à ce stade que le questionnaire initial du premier tour ne comportait ni la centralisation ni la décentralisation, mais ces modes de gouvernance découlent de la réponse des répondants suite au premier tour de l'enquête. Ceci représente un atout de la méthode DELPHI adoptée pour cette étude, qui encourage une discussion interactive sur les avis et les positions des répondants, ce qu'on n'aurait pas eu avec un sondage par exemple.* L'analyse des différentes réponses révèle que 3/7 sont en faveur d'une centralisation sous la responsabilité d'un CDO ; 2/7 sont en faveur de la centralisation et de la décentralisation selon le contexte et enfin 2/7 sont pour un organisme externe à l'équipe de traitement de données, peu importe si le mode est centralisé ou pas. Les uns et les autres justifient leur position via des arguments qui sont présentés en annexe 11 de ce document.

*Constat 4 découlant de la quatrième question du premier tour de l'enquête : **mécanisme de contrôle***

Afin de permettre la viabilité d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données, les experts ont été appelés à se prononcer sur les mécanismes de contrôle en entreprise. Deux propositions qui mettent en opposition un contrôle interne et externe des pratiques des scientifiques des données ressortent de l'analyse préliminaire des réponses de l'enquête du premier tour. Au second tour, ces deux propositions ont été soumises à l'ensemble des experts afin de recueillir leur avis. **(Une fois de plus un avantage de la méthodologie DELPHI, comme c'est le cas pour le constat 3.)**

Le quatrième constat a conduit à plusieurs questions. À la première qui consiste à savoir si les experts partagent l'idée de la création d'une commission nationale pour l'éthique de l'exploitation des données en intelligence d'affaires et analytique (contrôle externe), les résultats sont les suivants : 2/7 sont d'accord pour l'existence d'un tel organisme. 1/7 prône l'existence de deux modes de contrôle : paragouvernemental et un outil interne et automatisé. 2/7 sont en faveur d'un organisme paragouvernemental cependant, ils émettent une réticence quant à son fonctionnement effectif, car pour eux, les commissions sont par expérience des occasions de pertes de temps et d'argent. 1/7 n'est ni en faveur ni en défaveur, car pour lui, un tel mécanisme devrait davantage être orienté vers le respect d'un cadre réglementaire autour de la donnée et enfin 1/7 est complètement en désaccord avec l'idée.

La deuxième question du quatrième constat consiste à savoir comment fonctionnerait l'organisme dont il est question dans le paragraphe précédent. À cette question, les réponses sont une fois de plus variant et dépendent de l'avis des experts à la question précédente. Toutefois, ceux qui étaient favorables à l'existence d'un organisme paragouvernemental à la question précédente proposent les éléments de réponses suivants :

- S'inspirer du fonctionnement des CPA, ou encore qu'il fonctionne sensiblement de la même façon que la Commission de l'éthique en science et en technologie (CEST) en animant des réflexions et des débats

sur l'éthique en matière d'analyse et d'exploitation de données afin d'élaborer les normes.

- Existence un portail d'information accessible permettant de s'informer sur les lois, les jurisprudences, les contextes des absences de conformité ; puis l'existence d'un service de gestion des demandes d'information, et /ou de contrôle, de plaintes.
- L'existence des règles et des sanctions en cas de non-respect de celles-ci, accompagnée par des audits rigoureux des données et contrôles de l'application des règles à l'interne dans les compagnies.

La troisième question du quatrième constat consiste à savoir si les répondants sont en faveur de la mise en place d'un outil de contrôle interne à l'entreprise et automatisé. À cette question, tous les répondants (7/7) ont répondu a priori « OUI ». Cependant, 3/7 ont accompagné leur « Oui » de quelques réserves. En effet, le premier estime que ce ne sera pas facile de procéder à cette mise en place. Le deuxième pense que l'outil devrait être en place uniquement pour les services gouvernementaux et enfin le troisième est en faveur de cet outil uniquement dans un contrôle standard à l'aide d'une gouvernance des données, où l'application de politiques s'avère possible ; mais demeure complètement contre l'automatisation dans un cadre prospectif ou de recherche.

Toujours dans le cadre de contrôle, la cinquième question du quatrième constat consiste à savoir de la part des experts, comment fonctionnerait l'outil interne et automatisé dont il est question au paragraphe précédent. Une fois de plus, plusieurs avis ressortent ici : certains proposent que cela prenne la forme d'un support aux décisions reliées à l'éthique et déontologie, sans toutefois être une boîte noire qui prend des décisions sans aucune explication. D'autres prônent l'élaboration des règles et des contrôles via des audits permanents.

Enfin, toujours dans le cadre du mécanisme de contrôle, les experts ont été questionnés sur la cohabitation ou non des deux modes de contrôle : paragouvernemental et outil interne et automatisé de contrôle. La réponse des experts est unanime. En effet, tous (7/7) sont d'accord pour dire que si les deux existent, ils devront cohabiter, car disent-ils, ils sont complémentaires.

Synthèse des résultats du second tour de l'enquête DELPHI portant sur les éléments liés à l'exploitation éthique des données.

Les avis des experts au second de tour de l'enquête ont permis de noter de manière générale une convergence sur la plupart des éléments soumis à leur appréciation. En effet, de l'ensemble des opinions récoltées chez les experts lors du second tour de l'enquête DELPHI, ressortent les résultats suivants : 1) il y a un consensus autour des termes confidentialité, sécurité, exactitude, secret où au minimum 5/7 des répondants s'accordent sur la même compréhension et le même sens de chaque terme mentionné. Par contre les termes transparence et intégrité ne font pas toujours l'unanimité parmi les experts ; 2) Concernant la personne responsable de l'effectivité du code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique, les opinions sont plutôt divergentes. Ici, on a trois groupes d'avis : ceux qui sont en faveur d'une gouvernance centralisée auprès d'un CDO (*Chief Data Officer*), ceux qui sont favorables à une gouvernance décentralisée où chacun doit être responsable à son niveau et enfin ceux qui sont à la fois favorables à une gouvernance centralisée et à une gouvernance décentralisée, tout dépendant de la taille de l'entreprise. 3) Pour ce qui est du classement des termes liés à l'utilisation éthique des données, on remarque également une convergence qui reflète la compréhension du sens de chaque terme de la question 1. On peut remarquer ici qu'il y a un changement par rapport au classement du premier tour. 4) concernant le mécanisme de contrôle de l'effectivité de la mise en application du code d'éthique, les avis des experts sont variés. Certains sont en faveur de l'existence d'un organisme étatique, chargé de veiller à la mise en application et du respect du code d'éthique et de déontologie en entreprise ; d'autres sont pour un contrôle partagé entre un organisme paragouvernemental et l'entreprise et un dernier groupe n'est ni en faveur ni en défaveur, car pour eux, ils n'ont pas assez de connaissances sur la question pour pouvoir donner un avis éclairé. Cependant, les experts s'accordent unanimement au sujet de la mise en place d'un système de contrôle interne et automatisé permettant de surveiller la mise en application efficace du code d'éthique et de déontologie, mais tout en reconnaissant pour certains d'entre eux que ce ne sera pas une tâche facile. Aussi, certains mettent en garde sur le fait que cette solution ne soit pas une boîte noire qui prend des décisions sans fournir une explication.

b) Résultat des questionnaires d'équipes de scientifiques de données

Le deuxième profil d'experts qui a participé à l'enquête pour la présente étude est constitué de gestionnaires d'équipe d'intelligence d'affaires et analytique.

Dans les lignes suivantes sont présentées leurs réponses au premier et au second tour de l'enquête. Tout comme les scientifiques des données, les gestionnaires ont eu à répondre à cinq questions au premier tour. Les questions étaient identiques, mais ces derniers devaient répondre en tant que gestionnaire, c'est-à-dire n'étant pas impliqués au premier rang dans l'exploitation des données.

Question 1 du questionnaire du premier tour de l'enquête DELPHI portant sur les éléments liés à l'utilisation éthique des données.

- La première question consiste à savoir si les concepts suivants : l'anonymat, la confidentialité, la sécurité, le secret, l'exactitude, l'intégrité et la transparence, liés à l'utilisation éthique des données en IA&A tels que ressorti par une revue de littérature sur ce sujet sont tous considérés déterminant par les experts en exploitation de données. À cette question, **3/11 des répondants n'a pas fourni de réponse, 6/11 est totalement d'accord et 2/11 pensent que tout dépend de l'usage de la donnée.**
- La deuxième partie de cette première question consiste à donner un sens à chacun des concepts. Les réponses des experts sont les suivantes :
 - **Le concept Anonymat** est compris par tous les répondants, soit **11/11** comme le fait de ne pas être en mesure de lier la donnée à une personne spécifique.
 - **Le concept confidentialité** : **4/11** des répondants accordent le sens de **non-divulcation de la donnée au terme confidentialité. 3/11** quant à eux lui donnent le sens **de nonaccès ou restriction d'accès aux seules personnes autorisées. Et 4/11** enfin lui donne le sens de **garder secret l'information détenue et ne pas permettre que celle-ci permette d'identifier la personne référencée.**
 - **Le concept sécurité** : **7/11** des répondants associent au terme sécurité **l'absence d'accès environnemental aux données par des personnes**

non autorisées. Tandis **4/11** lui donnent le sens **de sécuriser les données en préservant son intégrité ainsi que tout danger qui pourrait mettre la donnée en danger.** Si un groupe met l'accent sur l'accès à la donnée, l'autre groupe quant à lui, met l'accent sur l'intégrité et la confidentialité de la donnée.

- **Le concept secret :** alors que **5/11** estiment que **le terme trouve son sens dans la combinaison de confidentialité, anonymat et sécurité.** **2/11** disent que **le sens n'est pas clair** et **4/11** estiment que **c'est le fait de ne pas divulguer ou dévoiler une information ou le but de l'analyse d'une donnée.**
- **Le concept exactitude :** **3/11** des répondants **donnent le sens de vérité, ou ce qui est conforme à la réalité** au terme exactitude. **5/11** lui donnent le sens **« d'exact » et démunie de toute falsification.** **1/11** associent le sens de **« bonnes » et d'utilisables.** **1/11** lui donnent le sens **de bonne qualité et la possibilité de s'y fier** et enfin **1/11** lui donnent le sens **de précision et de certitude de l'information dérivée de la donnée.**
- **Le concept intégrité :** **3/11** n'ont pas fourni de réponse. Alors que **1/11** dit qu'il s'agit de **la bonne intention de ceux qui traitent la donnée ;** **6/11** quant à eux estiment qu'il s'agit **pour la donnée de garder son état initial durant le processus de traitement et ne subir aucune altération.** Enfin pour **1/11** c'est le fait **d'intégrer la donnée dans un ensemble plus global avec possibilité que celle-ci soit validée en recoupant avec d'autres données références.**
- **Le concept transparence :** **5/11** des répondants lui donnent le sens **d'énoncer clairement l'utilisation finale de la donnée ainsi que la façon dont elles sont utilisées.** **1/11** l'associe plutôt au fait **de rendre publique la provenance de la donnée.** **2/11** lui donnent le sens de **rendre accessible la donnée elle-même et non la source lorsqu'elle concerne l'opinion publique.** **2/11** disent qu'il s'agit de **documenter de façon complète dans un langage non codé les processus relatifs à la protection, exploitation et diffusion de l'information.** Enfin **1/11** lui

associe le sens **de transparence au niveau de l'analyse faite avec les données.**

Question 2 du questionnaire du premier tour de l'enquête DELPHI portant sur les autres éléments déterminants dans l'élaboration d'un code d'éthique dédiée à l'intelligence d'affaires et l'analytique et qui ne faisaient pas partie de la liste proposée.

- **La deuxième question consiste à savoir quels sont les autres concepts que les répondants estiment qu'ils doivent être pris en compte dans un code éthique dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données.**

Les réponses du panel d'experts à cette seconde question concernant les autres concepts à considérer ont, dès le premier tour, révélé une divergence. Même si les concepts comme « gouvernance des données », « respect de la loi » reviennent à quelques reprises, on peut remarquer ici qu'il y a autant de concepts supplémentaires qu'il y a de répondants, chacun faisant une proposition selon sa compréhension de l'étude. À titre d'exemple, un expert propose d'intégrer la « Supervision par une partie indépendante » comme autre concept tandis qu'un autre propose « Stratégies : Art de coordonner des actions, de manœuvrer habilement pour atteindre un but » (voir les éléments proposés par les experts en annexe 12).

Question 3 du questionnaire du premier tour de l'enquête DELPHI portant sur le classement des concepts selon l'ordre de priorité du plus prioritaire au moins prioritaire.

- **La troisième question consiste à classer les termes selon un ordre de priorité.** Il faut dire qu'à cette question, les classements sont très différents de part et d'autre. Certains ont estimé que tous les termes étaient prioritaires au même niveau, d'autres ont fait des classements partiels selon l'importance qu'ils accordaient aux termes. Somme toute, un récapitulatif des résultats nous donne le classement suivant du plus prioritaire au moins prioritaire :
 - Transparence
 - Confidentialité

- Sécurité
- Anonymat
- Intégrité
- Exactitude
- Secret

Ces éléments ont été regroupés en trois grands groupes, selon **l'importance accordée** à chaque terme par les experts :

Premier groupe d'éléments prioritaires regroupe	Transparence (7/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en 1 ^{re} ou 2 ^e position), Confidentialité (5/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en 1 ^{re} ou 2 ^e position) et sécurité (4/11 l'ont mis en 1 ^{re} ou 2 ^e position)
Deuxième groupe d'éléments prioritaires regroupe	Anonymat (4/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en 1 ^{re} ou 2 ^e position) Intégrité (4/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en 2 ^e ou 3 ^e position)
Troisième groupe d'éléments prioritaires regroupe :	Exactitude (5/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en dernière position de leur classement) Secret (4/11 de ceux qui ont inclus ce concept dans leur classement l'ont mis en dernière position)

Question 4 du questionnaire du premier tour de l'enquête DELPHI portant sur la personne responsable de veiller à l'effectivité d'un tel code.

- **La quatrième question consiste à déterminer qui dans l'entreprise doit veiller à l'effectivité :**

Ici, une grande majorité s'accorde pour confier cette responsabilité au CDO (*Chief Data Officer* / Chef Officier des données (COD)), au CAO (*Chief Analytics Officer* / Chef Officier Analytique (COA)) au CDAO (*Chief Data*

Analytics Officer / Chef Officier de la Donnée et de l'Analytique (CODA)) ou à un haut cadre comme le VP marketing si la compagnie n'a pas un des trois postes précédemment cités. Cependant, certains ont pensé à la création d'un comité ou d'une équipe dédiée à l'éthique et déontologie en matière d'exploitation des données et le gestionnaire responsable de cette équipe ou de ce comité sera chargé de l'effectivité.

Question 5 du questionnaire du premier tour de l'enquête DELPHI portant sur le mécanisme de contrôle d'un tel code.

- **Enfin, la cinquième question consiste à savoir quels sont les mécanismes de contrôle de gestion :** plusieurs suggestions selon le concept sont proposées et qui relèvent en majorité de la gouvernance des données (audit) : Depuis le contrôle d'accès, en passant par le cryptage, l'utilisation limitée aux besoins, à la formation que la mise en place des fiches d'engagement que chaque professionnel doit signer et qui engage sa responsabilité.

Cependant, en plus de l'audit et d'autres mécanismes liés à la gouvernance des données, la mise en place d'un tableau bord avec indicateur permettant de surveiller chaque concept a été proposé.

Synthèse des résultats des gestionnaires au premier tour de l'enquête DELPHI portant sur les éléments liés à l'exploitation éthique des données.

De l'ensemble des opinions récoltées chez les experts lors du premier tour de l'enquête DELPHI, ressortent les résultats suivants : 1) il y a un consensus autour des termes anonymat et sécurité. En effet **tous** les experts (**11/11**) donnent le même sens au terme anonymat tandis **7/11** donnent le même sens au terme sécurité. 2) Concernant la personne responsable de l'effectivité du code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique, une grande majorité s'accorde pour un contrôle centralité auprès d'un haut cadre de l'entreprise comme CDO (*Chief Data Officer*), au CAO (*Chief Analytics Officer*) au CDAO (chef officier de la donnée et de l'analytique) ou le VP marketing. 3) pour ce qui est du classement des concepts considérés, on observe ici une divergence d'opinions, reflétant la compréhension que les experts ont des termes abordés à la question 1. 4) parlant du mécanisme de contrôle, selon une minorité des experts, il serait indiqué de penser à un outil sous forme de tableau de bord

avec indicateur permettant de surveiller chaque élément lié à l'utilisation éthique des données. D'autres suggestions ont été faites et relèvent plus de la gouvernance des données, il s'agit entre autres du contrôle d'accès, des limitations d'accès, du cryptage des données.

Deuxième tour (Round 2) : dépouillement, classification et consolidation des concepts liés à l'utilisation éthique des données.

À la suite du premier tour de l'enquête, des divergences ont été observées comme ce fut le cas également pour les scientifiques des données ; ce qui n'est pas tout à fait surprenant du fait même de la nature de l'enquête. Ces divergences ont été formulées également sous forme de constats et communiquées aux experts dans le but de souligner pour chaque question du premier tour la divergence observée et ce qui était attendu au second tour.

Ainsi dans les lignes suivantes, les constats qui ont animé le second tour de l'enquête sont présentés. Ce second tour s'est déroulé sur une période de 7 semaines, soit du 1^{er} avril avec l'envoi des questionnaires et s'est terminé le 22 mai 2019 avec le début de la compilation.

Un questionnaire basé sur quatre constats découlant du premier questionnaire a été remis à chacun des 11 participants du premier et de ces 11 participants, 7 ont retourné leurs formulaires complétés et il en ressort ce qui suit :

Constat 1 découlant de la première question du premier tour de l'enquête : Disparité dans la compréhension des termes liés à l'éthique

- Le premier constat portait sur la compréhension ou mieux le sens que les répondants donnaient aux différents éléments liés à l'éthique ; à ce niveau, il a été constaté une disparité dans la compréhension et il a été soumis à tous les répondants les différents sens que les uns et les autres donnaient aux termes afin qu'ils réagissent de nouveau suite à l'exposition des différentes compréhensions. On constate actuellement, une convergence à l'issue de ce second tour d'enquête. En effet, l'analyse révèle ceci :
 - **Le concept confidentialité** : 5/7 des répondants associent le sens de « non-accès ou restriction d'accès uniquement aux personnes

autorisées »; 1/7 des répondants pense qu'il s'agit plutôt de « garder secrète l'information détenue et éviter que celle-ci permette d'identifier la personne référencée » et un 1/7 estime plutôt que « Ceux qui ont la responsabilité de traiter la donnée ne doivent divulguer aucune information sur ce qu'ils auraient pu constater lors de l'analyse, ainsi que les données ».

- **Le concept sécurité :** 4/7 des répondants pensent qu'il s'agit de « sécuriser les données en préservant son intégrité ainsi que tout ce qui pourrait mettre la donnée en danger ». 1/7 quant à eux disent qu'il s'agit de : « sécuriser les données de manière à préserver leur intégrité et les mettre à l'abri de toute menace à celle-ci » ; 1/7 nuancent un petit peu leur compréhension en lui donnant comme sens « l'action de protéger les données contre les accès non autorisés et contre toute altération de la donnée ». Dans ces trois groupes de répondants, *l'idée de préserver l'intégrité de la donnée ressort dans leur compréhension du terme sécurité en plus de l'idée de protéger la donnée*. Un dernier groupe soit 1/7 met l'accent sur la protection pure et simple de la donnée ; pour ce groupe le concept sécurité a le sens de « mettre la donnée à l'abri de tout danger. »
- **Le concept secret :** 3/7 des répondants trouvent qu'il s'agit de la « combinaison de confidentialité, anonymat et sécurité ». 3/7 quant à eux estiment que son sens n'est pas **clair** ; et enfin 1/7 lui associe le fait de « ne pas divulguer ou dévoiler une information ou le but de l'analyse d'une donnée ».
- **Le concept exactitude :** 2/7 des répondants donnent le sens : « précision et certitude de l'information dérivée de la donnée ». 1/7 lui associe le sens de « précision et certitude de l'information ». 1/7 quant à lui l'associe à « bonne qualité et possibilité de s'y fier ». 1/7 également l'associe à ce qui « exact et démunie de toute falsification ». 1/7 dit qu'il s'agit de « vérité, conforme à la réalité ». 1/7 s'est abstenu de répondre. **On remarque ici une grande divergence dans la compréhension du sens de ce concept ; en effet, on a autant de compréhension que**

d'experts. Bien que les deux premiers groupes aient une compréhension presque similaire, il ressort tout même une légère différence, car le premier groupe parle de l'information dérivée de la donnée alors que le second se limite à l'information. Les autres groupes ont chacun leur compréhension du concept exactitude.

- **Le concept intégrité :** 7/7 des répondants estiment que c'est « Le fait que la donnée garde son état initial durant le processus de traitement et ne subisse aucune altération ».
- **Le concept transparence :** 3/7 des répondants lui donnent le sens d'« énoncer clairement l'utilité finale de la donnée ainsi que la façon dont elles sont utilisées »; 3/7 estiment plutôt qu'il s'agit de « documenter de façon complète dans un langage non codé les processus relatifs à la protection, exploitation et diffusion de l'information » ; enfin 1/7 quant à lui pense qu'il s'agit de « rendre accessible la donnée elle-même et non la source lorsqu'elle concerne l'opinion publique ».

*Constat 2 découlant de la deuxième question du premier tour de l'enquête et portant sur la classification des **Éléments prioritaires** dans la définition et la mise en œuvre d'un code d'éthique.*

- Le second constat lié à la classification des termes à la suite de la nouvelle compréhension faite par les répondants révèle qu'il y a eu un changement au niveau du classement. Le nouveau classement du plus prioritaire au moins prioritaire est le suivant :

2 ^e tour	1 ^{er} tour :
<ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité ○ Transparence ○ Intégrité ○ Anonymat ○ Exactitude ○ Secret 	<ul style="list-style-type: none"> ○ Transparence ○ Confidentialité ○ Sécurité ○ Anonymat ○ Intégrité ○ Exactitude ○ Secret

On remarque que la sécurité devient le premier élément le plus prioritaire en remplacement du concept de transparence dans le précédent classement, suivi de la confidentialité qui lui n'a pas changé ; tout comme exactitude et secret qui restent les deux éléments les moins prioritaires ; sans doute lié aussi au fait que leur compréhension est assez différente d'une personne à l'autre parmi les répondants ou bien le sens n'est pas tout à fait clair pour eux, comme c'est le cas pour le concept « secret ».

Ces éléments ont été regroupés en trois grands groupes, selon l'importance accordée à chaque terme :

Premier groupe d'éléments prioritaires regroupe	Confidentialité (5/7 experts l'ont mis en 1 ^{re} , 2 ^e ou 3 ^e position), sécurité (6/7 experts l'ont mis en 1 ^{re} , 2 ^e ou 3 ^e position)
Deuxième groupe d'éléments prioritaires regroupe	transparence (3/7 experts l'ont mis en 1 ^{re} position), Intégrité (2/7 experts l'ont mis en 2 ^e position),
Troisième groupe d'éléments prioritaires regroupe	Secret (4/7 experts l'ont mis en 7 ^e position), Anonymat (3/7 experts l'ont mis en 5 ^e , 6 ^e ou 7 ^e position) et exactitude (5/7 experts l'ont mis en 5 ^e ou 6 ^e position)

Constat 3 découlant de la troisième question du premier tour de l'enquête :
Gouvernance de l'éthique

- **Le troisième constat portait sur la gouvernance de l'éthique dans l'organisation.** En effet, au premier tour de l'enquête, les experts devaient proposer selon eux qui dans l'organisme doit être responsable de l'effectivité du fonctionnement du code éthique en cas d'existence de celui-ci. Les résultats obtenus se sont déclinés en deux modes de gouvernance, à savoir une responsabilité centralisée et une responsabilité décentralisée. Face à ce constat, lors du second tour, il a été demandé aux experts de se prononcer au sujet de la centralisation auprès d'une personne-ressource ou de la décentralisation de la gouvernance de l'éthique en intelligence d'affaires et analytique ; ceci afin de

chercher une convergence vers l'une ou l'autre des propositions que les experts avaient initialement proposées. L'analyse des différentes réponses révèle que : **5/7** des répondants sont en faveur de la **centralisation** ; **1/7** prône la **décentralisation** et **1/7** est en faveur des deux, tout dépendant de la taille de l'entreprise. Les uns et les autres justifient leur position via des arguments qui sont disponibles en annexe du présent document.

*Constat 4 découlant de la quatrième question du premier tour de l'enquête et portant sur le : **mécanisme de contrôle***

- Au premier tour de l'enquête, il avait été demandé aux experts comment ils envisageaient le contrôle de l'effectivité du code d'éthique destiné à l'intelligence d'affaires et analytique. Plusieurs propositions ont été faites et parmi elles, deux propositions à savoir la création d'un organisme paragouvernemental d'une part et la création d'un outil interne et automatisé d'autre part sont ressorties parmi les plus pertinentes. Dans le deuxième questionnaire, ces deux propositions ont été soumises à l'ensemble des répondants afin qu'ils se prononcent notamment au sujet de l'existence d'un organisme paragouvernemental. Il ressort de l'analyse des différentes réponses que **3/7** sont **d'accord** avec la création d'un organisme paragouvernemental responsable du contrôle de la pratique éthique de l'intelligence d'affaires et l'analytique ; **3/7** sont en **désaccord** avec l'idée et **1/7** se questionne si oui ou non un tel organisme devrait exister, car pour cet expert, il existe déjà la loi sur la protection des renseignements personnels et les documents électroniques (LPRPDE) ; les consommateurs ayant déjà l'opportunité de soumettre des cas problématiques au Commissariat à la protection de la vie privée du Canada.
- Pour ce qui est du fonctionnement de cet organisme paragouvernemental, ceux qui sont en faveur de sa mise en place suggèrent que :
 - Cette commission établit des principes directeurs clairs concernant l'utilisation acceptable des données et qu'elle ait la légitimité de les faire respecter. Pour eux, il n'est surtout pas question ici de se fier à l'autorégulation par des entreprises.
 - Cette commission s'inspire du CNESST ou encore, des Normes de la Publicité. Elle pourrait prendre la forme d'un organisme indépendant,

sans but lucratif, chargé d'administrer un code d'éthique et de déontologie en matière de stockage, d'analyse et d'exploitation de la donnée. Ici par contre on voit un volet d'**autorégulation** et un volet externe à l'industrie.

- Le fonctionnement de cette commission soit similaire à celui des CPA qui émettent des lois et règlements mondiaux de type « **IFRS** ». Ceux-ci sont vérifiés par des contrôles internes par le biais de vérificateurs indépendants. Ces vérificateurs valident leur application et émettent leurs opinions indépendantes et objectives. Cet avis force les organisations à se prendre en main et elles deviennent donc plus matures avec le temps.
 - Enfin, il est aussi suggéré de faire évoluer la loi sur la protection des renseignements personnels et les documents électroniques (LPRPDE) pour intégrer la notion éthique de l'intelligence artificielle.
- Parlant toujours du mécanisme de contrôle, les répondants ont été appelés à se prononcer au sujet de la mise en place d'un outil interne et automatisé. Sur ce point, **6/7** répondants sont **en faveur de l'automatisation** et **1/7** partiellement d'accord (uniquement pour les tâches routinières de contrôle) et partiellement en désaccord (pour tout ce qui est jugement du contenu ou du processus du code). Pour ce qui est du fonctionnement de cet outil, plusieurs avis ressortent ici. Certains proposent que cela prenne la forme d'un soutien aux décisions reliées à l'éthique où un comité définit les métriques à surveiller; d'autres prônent la mise en place de techniques d'apprentissage automatique en plus des humains/experts qui sont à même de juger si le code d'éthique et de déontologie est respecté ; d'autres encore suggèrent une fois de plus de s'inspirer de la pratique des CPA pour faciliter l'implantation d'une telle pratique et faire des audits sur une base annuelle pour s'assurer que l'éthique est respectée.
- Enfin les experts ont été questionnés sur la cohabitation ou non des deux modes de contrôle : paragouvernemental et outil interne et automatisé de contrôle. **6/7** sont en faveur de la cohabitation des deux systèmes et **1/7** est plus ou moins en faveur, car pour lui, un mécanisme paragouvernemental est souvent peu flexible et s'adapte rarement assez vite pour bien suivre l'évolution de la pratique.

Synthèse des résultats du second tour de l'enquête DELPHI portant sur les éléments liés à l'exploitation éthique des données.

Comme pour le cas des scientifiques des données, les avis des gestionnaires au second de tour de l'enquête ont permis de noter de manière générale une convergence sur la plupart des éléments soumis à leur appréciation. En effet, de l'ensemble des opinions récoltées chez les experts lors du second tour de l'enquête DELPHI, il ressort les résultats suivants : 1) il y a un consensus autour des termes confidentialité, sécurité, intégrité où au minimum **4/7** des répondants s'accordent sur la même compréhension et le même sens de chaque terme mentionné. Par contre, les termes transparence, exactitude et secret ne font pas toujours l'unanimité parmi les experts ; 2) Concernant la personne responsable de l'effectivité du code d'éthique dédié à l'intelligence d'affaires et l'analytique, les experts sont favorables en grande majorité à la centralisation du rôle de responsable de l'effectivité. En effet **5/7** experts sont de cet avis, un expert (**1/7**) étant pour la centralisation et la décentralisation en fonction de la taille de la compagnie et un dernier expert (**1/7**) étant plutôt en faveur de la décentralisation quelque que soit la taille de l'entreprise. 3) Pour ce qui est du classement des termes liés à l'utilisation éthique des données, on observe également une convergence qui reflète la compréhension du sens de chaque terme de la question 1. On peut remarquer ici qu'il y a un changement par rapport au classement du premier tour **où le concept transparence passe du 1^{er} élément le plus prioritaire et se positionne comme 3^e élément le plus prioritaire et le concept sécurité qui au contraire passe du 3^e élément le plus prioritaire au 1^{er} élément le plus prioritaire après le second tour d'enquête**. 4) parlant à présent du mécanisme de contrôle de l'effectivité de la mise en application du code d'éthique, les avis des experts ont été requis au sujet d'un organisme paragouvernemental d'une part et de la création d'un outil interne et automatisé d'autre part. Alors que les avis sont partagés au sujet de l'existence d'un organisme paragouvernemental, les experts sont en grande majorité (**6/7**) d'accord avec l'idée de la mise en place d'un outil interne et automatisé en vue du contrôle de la mise en application efficace du code d'éthique et de déontologie. À la question de savoir si les deux outils pouvaient cohabiter, les experts s'accordent tous (**7/7**) pour dire que si les deux existent, alors, ils doivent cohabiter.

Cette section consistait à décrire les résultats de l'enquête menée pour le compte de cette recherche. Il a été premièrement présenté les résultats des scientifiques des

données au premier et au second tour d'enquête ; puis une synthèse des résultats de chaque tour a été faite. Ensuite, le même scénario a été répété pour les gestionnaires d'équipes de scientifiques des données. La présentation a été faite de manière séparée afin de ressortir les éléments propres à chaque population constituant les répondants de cette enquête. Dans la section suivante, il sera question de discuter des résultats obtenus en balisant l'interprétation et en signifiant les limites observées.

2. Discussion

Après avoir présenté les résultats de la présente enquête dans la section précédente, il est à présent question dans cette section de la mise en perspective de ces résultats. Premièrement, il est question de comparer les résultats fournis par les deux groupes de répondants. Deuxièmement, examiner ces résultats au regard de la littérature scientifique et enfin ressortir quelques conclusions importantes découlant de cette recherche.

a) Comparaison des résultats proposés par les deux groupes des répondants au premier tour

L'enquête soutenant la présente étude cible deux catégories de répondants, le premier constitué de scientifiques des données directement impliqué dans l'exploitation quotidienne des données et la deuxième catégorie, les gestionnaires d'équipes de scientifiques des données qui eux, ne sont pas directement impliqués dans l'exploitation quotidienne des données. Le but de ce choix est de voir comment les deux groupes conçoivent l'exploitation éthique des données. Ont-ils la même perception des enjeux en vue ? Ont-ils les mêmes priorités vis-à-vis à une exploitation éthique des données ? De ce fait au premier tour de l'enquête, les deux groupes ont reçu exactement les mêmes questions, mais chacun devait répondre en fonction de son rôle vis-à-vis de la donnée. Dans le tableau suivant, on compare les résultats finaux en se limitant à l'avis de la majorité des répondants pour chaque groupe, sauf si les réponses d'une minorité nécessitent d'être comparé. Le résultat comparatif est le suivant :

NB : la première question du tableau 4 est évaluée par rapport aux 11 experts de chacune des deux populations ayant participé au premier tour. Toutes les autres questions sont évaluées par rapport aux 7 experts de chacune des populations ayant participé au deuxième tour de l'enquête.

Tableau 4

Résultat comparatif des scientifiques de données et des gestionnaires

scientifiques de données	Gestionnaires
L'anonymat, la confidentialité, la sécurité, le secret, l'exactitude, l'intégrité et la transparence sont-ils tous des concepts déterminants pour un code d'éthique dédié à l'intelligence d'affaires et analytiques?	
<ul style="list-style-type: none"> - 3/11 ont dit « oui » - 2/11 ont dit « ça dépend » - et 6/11 n'ont pas répondu 	<ul style="list-style-type: none"> - 6/11 ont dit « oui », - 2/11 ont dit « ça dépend » - et 3/11 n'ont pas répondu
Sens ou compréhension des concepts	
<ul style="list-style-type: none"> - Anonymat : tous les répondants s'accordent autour de « l'idée de ne pas dévoiler l'identité d'une personne en rapport avec un fait » - Confidentialité : 7/7 des répondants s'accordent autour du sens « la donnée est vue et utilisée uniquement par des personnes autorisées. ». - Sécurité : 6/7 des répondants pensent que c'est « l'action de protéger les données contre les accès non autorisés et contre toute altération de la donnée. » - Secret : 5/7 des répondants pensent qu'il s'agit de : « garder une discrétion dans le but de garantir la confidentialité ». 	<ul style="list-style-type: none"> - Anonymat tous les répondants s'accordent de « ne pas être en mesure de lier la donnée à une personne spécifique. » - Confidentialité : 5/7 des répondants associent le sens de « nonaccès ou restriction d'accès uniquement aux personnes autorisées ». - Sécurité : 4/7 des répondants pensent qu'il s'agit de « sécuriser les données en préservant son intégrité ainsi que tout ce qui pourrait mettre la donnée en danger donnée. » - Secret : 3/7 des répondants trouvent qu'il s'agit de la « combinaison de confidentialité, anonymat et sécurité ». - Exactitude : 2/7 des répondants

<ul style="list-style-type: none"> - Exactitude : 5/7 des répondants disent qu'il s'agit de la « représentation fidèle de ce qu'elle est censée effectivement représenter ». - Intégrité : 3/7 des répondants lui donnent le sens de « <i>donnée complète et fiable de bout en bout ainsi que l'intégrité des professionnels des données</i> ». - Transparence : 4/7 lui donnent comme sens le fait « <i>d'expliquer la méthodologie et ne pas obstruer la vue d'ensemble des pairs lors du partage de la méthodologie et des résultats.</i> » 	<p>donnent le sens : « précision et certitude de l'information dérivée de la donnée ».</p> <ul style="list-style-type: none"> - Intégrité : 7/7 des répondants estiment que c'est « Le fait que la donnée garde son état initial durant le processus de traitement et ne subisse aucune altération ». - Transparence : 3/7 des répondants lui donnent le sens d'« énoncer clairement l'utilité finale de la donnée ainsi que la façon dont elles sont utilisées » ; 3/7 estiment plutôt qu'il s'agit de « documenter de façon complète dans un langage non codé les processus relatifs à la protection, exploitation et diffusion de l'information. »
Classement des concepts selon l'ordre de priorité allant du plus prioritaire au moins prioritaire.	
<ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité ○ Transparence ○ Exactitude ○ Intégrité ○ Anonymat ○ Secret <p>En termes de regroupement, on a :</p> <p>1^{er} groupe :</p> <ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité 	<ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité ○ Transparence ○ Intégrité ○ Anonymat ○ Exactitude ○ Secret <p>En termes de regroupement, on a :</p> <p>1^{er} groupe :</p> <ul style="list-style-type: none"> ○ Sécurité ○ Confidentialité

<p>2^e groupe :</p> <ul style="list-style-type: none"> ○ Transparence ○ Exactitude ○ Intégrité <p>3^e groupe :</p> <ul style="list-style-type: none"> ○ Anonymat ○ Secret 	<p>2^e groupe :</p> <ul style="list-style-type: none"> ○ Transparence ○ Intégrité <p>3^e groupe :</p> <ul style="list-style-type: none"> ○ Anonymat ○ Exactitude ○ Secret
<p>Qui dans l'organisation devrait être responsable de veiller à l'effectivité c.-à-d. à la mise en opération effective, mesurée et pérenne des éléments liés à l'utilisation éthique des données</p>	
<p>3/7 sont en faveur d'une centralisation sous la responsabilité d'un CDO ; 2/7 sont en faveur de la centralisation et de la décentralisation selon le contexte et enfin 2/7 sont pour un organisme externe à l'équipe de traitement de données, peu importe si le mode est centralisé ou pas.</p>	<p>5/7 des répondants sont en faveur de la centralisation auprès d'un haut cadre comme CDO, au CAO, au CDAO ou le VP marketing. 1/7 expert est pour la centralisation et la décentralisation en fonction de la taille de la compagnie et 1/7 expert en faveur de la décentralisation, quel que soit la taille de l'entreprise.</p>
<p>Quels sont les mécanismes de contrôle de gestion à considérer pour chacun des éléments retenus</p>	
<p>- Plusieurs propositions qui relèvent davantage de la gouvernance (ce qui situe à l'étape 4 de la figure 4 : modèle de maturité de Gartner, 2008) des données plutôt que du contrôle en soi sont proposées. Cependant, une minorité propose la création d'une commission nationale de norme éthique en matière d'analyse et d'exploitations des données</p>	<p>- Plusieurs suggestions selon le concept sont proposées et relèvent en majorité de la gouvernance des données (audit) : Depuis le contrôle d'accès, en passant par le cryptage, l'utilisation limitée aux besoins, à la formation que la mise en place des fiches d'engagement que chaque professionnel doit signer et qui engage sa responsabilité. Cependant,</p>

et une autre minorité a proposé un mécanisme plutôt général en deux approches statiques et dynamiques (tableau de bord permettant de surveiller les concepts considérés).	en plus de l'audit et autres mécanismes liés à la gouvernance des données, la mise en place d'un tableau bord avec indicateur permettant de surveiller chaque concept a été proposé.
Suite aux réponses de la question précédente, les experts ont été questionnés sur l'idée de la création d'une commission de norme éthique et voici les réponses :	
- 2/7 sont d'accord pour l'existence d'un tel organisme. 2/7 sont en faveur d'un organisme paragouvernemental cependant, ils émettent une réticence quant à son fonctionnement effectif, car pour eux, les commissions sont par expérience des occasions de pertes de temps et d'argent. 1/7 n'est ni en faveur ni en défaveur, car pour lui, il ne possède pas assez d'informations sur le fonctionnement de tels organismes lui permettant de donner un avis éclairé et enfin 1/7 est complètement en désaccord avec l'idée et 1/7 prône l'existence de deux modes de contrôle : paragouvernemental et un outil interne et automatisé.	- 3/7 sont d'accord avec la création d'un organisme paragouvernemental responsable du contrôle de la pratique de l'IA&A de manière éthique ; 3/7 sont en désaccord avec l'idée et 1/7 se questionne si oui ou non un tel organisme devrait exister, car pour cet expert, il existe déjà la loi sur la protection des renseignements personnels et les documents électroniques (LPRPDE).
Comment envisager le fonctionnement de la commission de norme éthique en matière d'analyse et d'exploitation des données ?	
- S'inspirer du fonctionnement des CPA, ou encore qu'il fonctionne sensiblement de la même façon que la Commission de l'éthique en science et en technologie (CEST) en animant des réflexions et des débats sur l'éthique en	- Cette commission établit des principes directeurs clairs concernant l'utilisation acceptable des données et qu'elle ait la légitimité de les faire respecter. Il n'est surtout pas question ici de se fier à l'autorégulation par des entreprises.

<p>matière d'analyse et d'exploitation de données afin d'élaborer les normes.</p> <ul style="list-style-type: none"> - Existence d'un portail d'informations accessible permettant de s'informer sur les lois, les jurisprudences, les contextes des absences de conformité ; puis l'existence d'un service de gestion des demandes d'information, et /ou de contrôle, de plaintes. - L'existence des règles et des sanctions en cas de non-respect de celles-ci, accompagnée par des audits rigoureux des données et contrôles de l'application des règles internes dans les compagnies. 	<ul style="list-style-type: none"> - Ou bien s'inspirer du CNESST ou encore, des Normes de la Publicité. Elle pourrait prendre la forme d'un organisme indépendant, sans but lucratif, chargé d'administrer un code d'éthique en matière de stockage, d'analyse et d'exploitation de la donnée. Ici par contre on voit un volet d'autorégulation et un volet externe à l'industrie. - Le fonctionnement de cette commission soit similaire à celui des CPA qui émettent des lois et règlements mondiaux de type «IFRS». Ceux-ci sont vérifiés par des contrôles internes par le biais de vérificateurs indépendants. Ces vérificateurs valident leur application et émettent des opinions indépendantes et objectives. Cet avis force les organisations à se prendre en main et elles deviennent donc plus mature avec le temps.
<p>les experts ont été questionnés sur l'existence d'un outil de contrôle interne et automatisé et voici les réponses :</p>	
<p>Tous les répondants (7/7) ont répondu a priori « OUI ». Cependant, 3/7 ont émis des réserves à leur « Oui ». En effet, le premier estime que ce ne sera pas facile de procéder à cette mise en place. le deuxième pense que l'outil devrait être en place uniquement pour les services</p>	<p>- 6/7 répondants sont en faveur de l'automatisation et 1/7 partiellement d'accord (uniquement pour les tâches routinières de contrôle) et partiellement en désaccord (pour tout ce qui est jugement du contenu ou du processus du code.)</p>

gouvernementaux et enfin le troisième est en faveur de cet outil uniquement dans un contrôle standard à l'aide d'une gouvernance des données, où l'application de politiques s'avère possible cependant cette même personne est complètement contre l'automatisation dans un cadre prospectif ou de recherche	
Les experts ont été questionnés sur la cohabitation des modes de contrôles à savoir un outil interne automatisé et un organisme externe	
Tous (7/7) ont répondu « OUI » si les deux existent, ils doivent cohabiter	Tous (7/7) ont répondu « OUI » si les deux existent, ils doivent cohabiter

Il ressort du tableau précédent que les scientifiques de données et les gestionnaires ont en majorité des réponses similaires aux différentes questions qui leur ont été soumises. Cependant, quelques spécificités sont à noter.

À la question de savoir si l'anonymat, la confidentialité, la sécurité, le secret, l'exactitude, l'intégrité et la transparence sont tous des concepts déterminants pour un code d'éthique dédié à l'intelligence d'affaires et analytique. Une grande majorité des gestionnaires a répondu « oui » tandis qu'une grande majorité des scientifiques des données n'a pas répondu à la question. **On pourrait se demander si l'absence de réponse de la majorité des scientifiques des données est due à un désintérêt ou c'est simplement parce qu'ils ne sont pas suffisamment instruits sur les concepts pour pouvoir donner une réponse éclairée.** Dans un cas comme dans un autre, ceci a des implications sur l'attitude que ces derniers peuvent adopter vis-à-vis de la donnée. Il est naturel de faire attention lorsqu'on connaît l'importance et la conséquence de tout ce qui a rapport avec l'exercice d'une fonction. Par contre, on peut être négligeant si on n'est pas au fait de ces éléments. Il est donc nécessaire d'intégrer l'aspect formation sur les éléments en question auprès de ce groupe (scientifique de données) d'une part et de vulgariser l'impact de ces éléments dans le processus de traitement de données, dans le processus de prise de décision et sur l'image de marque de l'organisation ainsi que leur

importance et leurs enjeux dans une perspective d'exploitation éthique des données d'autre part.

Pour ce qui est du sens de chaque concept, on peut remarquer :

- À l'intérieur de chaque groupe de répondants, la majorité s'accorde sur la plupart des concepts. **Cependant, le fait d'être unanime n'est pas naturel, mais s'observe plutôt à la suite d'une itération où les experts sont exposés aux avis les uns des autres.** On peut en déduire, qu'à l'intérieur d'une organisation, on gagnerait à exposer les professionnels comme les cadres aux éléments liés à l'exploitation éthique des données dans un contexte de groupe et non de manière individuelle.
- Quand on compare le sens accordé aux concepts par les scientifiques des données à ceux accordés par les gestionnaires, on réalise que les deux groupes ont la même compréhension des concepts anonymat, confidentialité et sécurité. Pour les autres concepts, on observe des différences parfois subtiles, mais qui peuvent avoir des implications et/ou des conséquences dans l'utilisation éthique des données. Parfois ces différences sont importantes. Serait-elle due au rôle que chacun d'entre eux a vis-à-vis de la donnée ?

Comme illustration d'une des **différences subtiles**, le concept exactitude. En effet, pour le concept **exactitude** les scientifiques de données s'accordent en majorité autour de la définition suivante : « **représentation fidèle de ce qu'elle est supposée effectivement représenter** » tandis que pour les gestionnaires, il s'agit de la « **précision et certitude de l'information dérivée de la donnée** ». En vérifiant les synonymes du mot « exactitude », on se rend compte que « fidélité » employé par les scientifiques de données à travers le mot « fidèle » et « précision » employé par les gestionnaires font tous les deux parties de la liste des synonymes de « exactitude », on pourrait a priori dire que les deux groupes ont la même compréhension du terme. Par contre, lorsqu'on regarde de près leur compréhension, on voit qu'il y a une différence subtile dans leur compréhension en ceci que, pour les scientifiques de données, exactitude veut dire que la donnée représente fidèlement ce qu'elle est supposée effectivement représenter ; ceci sous-entend qu'on peut établir une relation un à un (nommée exactitude) entre la

donnée et ce qu'elle représente. Pour les gestionnaires, exactitude veut dire précision et certitude de l'information dérivée de la donnée ; ici on sous-entend que le concept n'est pas lié à la donnée brute en soi, mais à l'information que la donnée véhicule. Sans faire une étude étymologique de « donnée » et « information », on voit qu'il s'agit de deux concepts différents, bien qu'étroitement liés dans le contexte de l'intelligence d'affaires et analytique. **Les gestionnaires sont dans une perspective de la finalité décisionnelle, tandis que les scientifiques de données sont dans une perspective de l'exploitation si on s'en tient à la définition même de l'intelligence d'affaires et l'analytique.**

Comme illustration d'une des **différences importantes** : le concept **secret**. En effet, pour les scientifiques de données, il s'agit de : « *garder une discrétion dans le but de garantir la confidentialité.* » Tandis que pour les gestionnaires, il s'agit de la « combinaison de confidentialité, anonymat et sécurité ». **Doit-on sous-entendre qu'en assurant à l'effectivité des éléments de sécurité, d'anonymat et de confidentialité on garantit certainement le secret?** Répondre à cette question pourrait faire l'objet d'une étude étant donné l'envergure et la pertinence dans un contexte d'exploitation de données. Toutefois, la différence est apparente et peut être débattue.

Parlant du classement des concepts, on a pratiquement le même classement de part et d'autre. En effet, les trois premiers éléments ainsi que le dernier élément sont les mêmes dans les deux classements. Par contre, les deux groupes n'accordent pas exactement la même importance à chaque élément. C'est la raison pour laquelle, lorsqu'on regroupe les éléments du classement, on observe une différence entre les différents groupes. Cette différence vient renforcer l'observation faite plus haut concernant la compréhension et le sens que les uns et les autres donnent aux différents concepts, notamment le cas de l'« exactitude ». On a signifié une différence subtile au niveau de la compréhension, mais qui pouvait avoir des conséquences importantes, l'une de ces conséquences réside ici dans la priorité accordée à l'élément. En effet, pour les scientifiques de données, l'exactitude est le 4^e élément plus prioritaire lié à l'exploitation éthique des données, après la sécurité, la confidentialité et la transparence ; tandis que pour les gestionnaires, c'est l'avant-dernier élément moins

prioritaire juste avant le concept secret. **Ceci renseigne sur le fait que les scientifiques des données ont conscience de la nécessité et de l'importance que les données puissent représenter fidèlement ce qu'elle est censée représenter et sans doute en tiennent compte dans l'exercice de leur fonction au quotidien.** Dans la revue de littérature, on a souligné qu'il n'était pas éthique de faire mentir les données et le livre *How To Lie With Statics* (Huff, 1954) a été évoqué à ce sujet ; même si dans le contexte on parlait de biaiser intentionnellement les décisions ; si les données ne sont pas exactes, elles peuvent avoir les mêmes conséquences, celles de biaiser une décision stratégique.

Pour ce qui est de la personne responsable de l'effectivité du code éthique dans l'organisation, les gestionnaires sont en majorité favorables pour une centralisation auprès d'un haut responsable de la compagnie tandis que les avis sont partagés parmi les scientifiques des données. Ceci sous-entend qu'au niveau de la maturité de l'intelligence d'affaires et l'analytique, les scientifiques de données ont un degré de maturité faible quant à ce qui a trait à l'étape de la gouvernance et la gestion des processus. En effet, dans le modèle de maturité de Gartner (2008) présenté par la figure 4, les notions de gouvernance et le rôle d'un CAO ressortent lorsqu'on a atteint le niveau 4/5 de maturité en intelligence d'affaires. On peut donc dire que sur cette question, les gestionnaires ont une longueur d'avance sur les scientifiques de données. De plus, la dimension gouvernance du modèle de maturité IAM (figure5) indique clairement que le mode de gouvernance décentralisé est caractéristique d'un niveau de maturité initial c'est-à-dire au premier stade de maturité. On observe donc que les scientifiques de données sont en majorité à ce stade initial pour ce qui est de la gouvernance de l'éthique. Une fois de plus, on pourrait se demander si ce manque de maturité n'est pas lié au volet formation. Cependant, on note que c'est un des scientifiques des données qui a émis l'idée d'un organisme étatique chargé du respect du code en industrie ; ce qui est intéressant et laisse entrevoir chez ce dernier premièrement une maturité dans sa compréhension de la gouvernance de l'éthique ; deuxièmement le souci de voir un œil externe contrôler la pratique de l'intelligence d'affaires et l'analytique et ainsi donner plus de crédibilité à l'activité au quotidien. Cette proposition rejoint aussi celle des autorités fédérales des États-Unis qui lors du procès contre le groupe Facebook dans l'affaire Cambridge Analytica, a décidé entre

autres de la mise en place d'un comité indépendant sur la protection de la vie privée, chargé de surveiller les activités de ce géant de communication dans le domaine des réseaux sociaux et de génération de contenus (Radio-Canada).

Concernant le mécanisme de contrôle du code, les positions adoptées par les deux groupes ont convergé après la première itération et sont passées d'un contrôle qui relève plus de la gouvernance que du contrôle de données en soi vers un mécanisme de contrôle interne et dynamique à l'aide d'un outil automatisé avec indicateur permettant de surveiller chacun des concepts. Les avis sont cependant partagés dans les deux groupes, quant à l'existence d'un organisme étatique de contrôle externe. Une moitié des gestionnaires est en faveur d'un organisme externe chargé du contrôle en plus du contrôle interne où il y a unanimité ; chez les scientifiques de données, 2/7 sont en faveur de l'existence d'un tel organisme, et le reste (5/7) est divisé entre deux positions à savoir : a) complètement en désaccord ; b) ne pas avoir assez d'information pour se prononcer.

À ce stade, une prochaine étape pourrait être de penser à la faisabilité de l'outil de contrôle interne de surveillance ; s'agira-t-il d'un outil d'apprentissage machine ? Ou plutôt un tableau de bord sur la base du rapport d'expertise lié aux cas de non-respect d'éthique décelé sur une période ? En s'inspirant du modèle de maturité IAM (Figure5), on remarque qu'en convergeant vers l'idée d'un outil de contrôle dynamique et interne à l'entreprise, les répondants se positionnent plus ou moins au niveau **3/5** pour ce qui est de la dimension processus. Ce qui est plutôt intéressant considérant que pour l'instant on est au balbutiement de l'existence du code d'éthique et de déontologie.

Enfin pour la cohabitation d'un outil de contrôle externe et d'un autre interne et automatisé, les deux groupes ont répondu en majorité être en faveur de la cohabitation.

En guise de conclusion à cette section, les avis des gestionnaires et des scientifiques de données au sujet des éléments liés à l'utilisation éthique des données révèlent que les uns et les autres ont **des compréhensions semblables sur certains éléments, mais différentes sur d'autres**. Ces différences peuvent avoir des conséquences dans la pratique du métier. Comme mentionné dans la section, **l'unanimité au niveau de certains avis dans l'un comme dans l'autre groupe n'est pas naturelle, mais a été**

observée après une itération où les experts étaient exposés à l'avis des autres membres de leurs groupes respectifs. La section suivante va examiner certains éléments de réponses des experts, à la lumière de la littérature.

a) Analyse des réponses par rapport à la littérature scientifique

Dans les sections précédentes, on a exposé les réponses qu'ont fournies les gestionnaires et les scientifiques de données par rapport aux différentes questions soumises à leur appréciation. On a ensuite comparé les résultats des deux groupes et relevé les éléments communs et les éléments différents. Dans cette section, on va analyser certains éléments de réponse à la lumière de la littérature, afin de voir où se situent les répondants par rapport à la littérature.

On a vu dans la section revue de littérature quelques enjeux éthiques liés à l'exploitation des données. Des définitions théoriques selon certains auteurs y ont été également citées. Durant l'enquête, les experts devaient se prononcer au sujet de leur compréhension de certains de ces enjeux. Dans les lignes suivantes, on va faire un état de leur compréhension versus la littérature énoncée.

En effet, pour le concept **anonymat**, on constate que Thomann et Wells (2013) et une grande majorité d'experts s'accordent pour dire que cet élément est important dans un code d'éthique dédié à l'intelligence d'affaires et analytiques. Cette affirmation est également constatée au niveau de l'intelligence artificielle. À cet effet, Saltz et Dewar (2019) le mentionnent comme l'une des trois importantes considérations éthiques des données massives. Les experts s'accordent aussi pour le sens du concept. Il est cependant nécessaire de considérer l'avis d'une minorité des répondants qui premièrement pense que l'importance du terme anonymat dépend si on a affaire aux données privées ou publiques. En cas de données publiques, il n'est pas nécessaire de considérer le terme comme important. Deuxièmement, cette même minorité attire l'attention sur la capacité de briser l'anonymat grâce aux analyses croisées d'un volume important de données et de ce fait reste retissant au sujet de l'importance du concept dans un code d'éthique et de déontologie dédié à l'exploitation éthique des données. À ce sujet, Saltz et Dewar (2019) affirment que 87% de personnes peuvent être identifiées avec précision suite à l'analyse croisée d'un jeu de données anonymes avec d'autres

sources grâce au code postal, sexe, année de naissance. Cette réserve pousse le chercheur à relever qu'il importe d'attaquer la question de l'anonymat dans un sens plus global toutefois en veillant à traiter les spécificités au cas par cas et selon le contexte.

Pour ce qui est du concept **confidentialité**, Thomann et Wells (2013) ainsi que la grande majorité des experts s'accordent sur le sens de terme. Ici, on peut remarquer que le sens que lui donnent les experts après une itération de l'enquête converge vers celui qu'on retrouve dans la littérature.

Parlant **de la sécurité**, pour Thomann et Wells (2013), tout comme pour la majorité des experts, le concept de sécurité renvoie à la notion de protection. Si les experts se sont limités à la protection des données, la littérature quant à elle voit la protection des données et des systèmes. Comme c'est le cas pour les deux premiers concepts, le concept sécurité apparaît aussi comme un enjeu éthique en intelligence artificielle. Vallejos et al. (2019) le soulignent dans le contexte des données en ligne et proposent par la même occasion des suggestions pour garantir la sécurité des données dans ce contexte.

Concernant le concept **secret**, au niveau des experts, on remarque une divergence entre les deux groupes de l'étude et entre chaque groupe et la littérature existence. Il a été mentionné dans la revue de littérature que les auteurs abordaient le concept sous divers angles. En effet, alors que Ming et al. (2015) en donnent une définition précise selon leur étude, McBride (2014) se questionne jusqu'à quel point on peut le garantir étant donné la liberté d'accès à l'information pour tous. Bien que ces divergences d'opinions soient observées, le concept représente un enjeu en exploitation des données comme l'indiquent Stark et Hoffmann (2019), Saltz et Dewar (2019) ainsi que Lanier et al. (2019) dans le domaine de l'intelligence artificielle et l'apprentissage machine. Il convient donc au niveau de l'intelligence d'affaires et l'analytique, de revoir ce concept dans un cadre académique/théorique, tout en gardant le contexte d'exploitation de données afin de s'assurer que toutes les parties prenantes : professionnels et académiciens s'accordent sur son sens et par conséquent sur les moyens de garantir une exploitation éthique des données en respectant cet élément.

Pour ce qui est de **l'exactitude**, une fois de plus, on observe une divergence entre les deux groupes et entre chaque groupe et la littérature existante. On remarque qu'au niveau de la littérature, on n'a pas beaucoup de référence sur ce concept. Cependant, dans le contexte du « *Machine Learning* », Lanier et al (2019) le mentionnent comme une préoccupation éthique à considérer. Ces derniers proposent par la même occasion des questions à se poser pour s'assurer la prise en compte de ce concept lors de l'analyse des données. Il convient donc de souligner la pauvreté de la littérature à ce niveau et espérer que ce manque sera comblé dans un futur proche par d'autres études sur le sujet afin de mieux outiller les professionnels et les académiciens.

Parlant du concept **transparence**, tout comme les deux concepts précédents, le concept transparence ne fait pas l'unanimité parmi les deux populations impliquées dans cette étude. On peut cependant remarquer qu'au niveau de la littérature, le sens tel que donné par Thomann et Wells (2013) est clairement énoncé de manière à faciliter la compréhension dans une perspective de mise en application. Dans le contexte de l'apprentissage machine et de l'intelligence artificielle, Lanier et al. (2019) mettent un accent sur la considération éthique de ce concept et proposent une série de questions qui permettront aux professionnels de s'assurer du traitement éthique des données vues sous cet angle. Il convient une fois de plus au niveau de l'intelligence d'affaires et l'analytique de divulguer ce concept au milieu des professionnels afin d'accroître leur compréhension et contribuer à l'intégration de ce concept vu sous l'angle de la littérature dans leur pratique quotidienne.

b) Conclusions importantes

Après avoir présenté les résultats d'enquête relative à la présente étude, il convient de noter quelques conclusions importantes. En effet, on retient de cette étude que :

- Les éléments que la littérature considère comme importants pour une exploitation éthique des données et qui doivent être pris en compte dans un code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique dans

ce contexte-là sont également perçus comme tels par une grande majorité des professionnels.

- Au niveau de la compréhension des éléments à considérer, un travail de formation d'une part et de vulgarisation d'autre part est à faire au niveau des professionnels tant pour les scientifiques de données que pour les gestionnaires.
- Toujours au sujet de la compréhension des éléments à considérer, il convient qu'au niveau académique, plus de travaux soient faits pour bien définir certains termes notamment le secret, la confidentialité, l'exactitude et veiller à leur vulgarisation.
- En ce qui a trait à la maturité de l'intelligence d'affaires et l'analytique, on peut voir sur la base du modèle de maturité de Gartner (2008) que les gestionnaires (par conséquent leur entreprise) se positionnent plus au niveau 4 dans la mesure où ils sont majoritairement dans une perspective de centralisation du rôle de responsable d'effectivité de l'éthique par rapport aux scientifiques de données où c'est une minorité qui entrevoit la centralisation de ce rôle.

Il est intéressant de voir qu'une grande majorité des experts des deux populations est en faveur d'un outil de contrôle interne et automatisé de l'effectivité du code d'éthique. Bien que cela ne soit pas arrivé de façon naturelle, mais plutôt après une itération, on peut en déduire le désir de voir chez les professionnels l'effectivité d'une gestion éthique dans l'organisation. Aussi le fait de suggérer un modèle de contrôle externe dont le fonctionnement est similaire à celui des CPA met en évidence le fait que les professionnels reconnaissent que la discipline est jeune et qu'elle peut s'inspirer d'un modèle établi et éprouvé.

3. Contributions

Cette section traite les contributions de la présente étude. Il est question ici de souligner premièrement ce que l'étude apporte sur le plan pratique. Comment peut-elle aider des organisations à construire un code d'éthique et de déontologie dédié à l'exploitation des données dans un contexte d'intelligence d'affaires et analytique ?

Deuxièmement, il est question de souligner la contribution théorique à travers le savoir et les informations que l'étude apporte au niveau de la recherche académique.

a) Contributions pratiques

En termes de contribution sur le plan pratique, on note que :

- Cette étude permet de valider quels sont les éléments importants à considérer pour la définition d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données.
- Elle permet de voir qu'à l'intérieure de l'organisation, les gestionnaires gagneraient à soumettre les scientifiques de données à des réflexions en groupe et non de manière isolée puisque la convergence sur les questions qui leur étaient exposées ne s'observait pas de façon naturelle, mais plutôt lorsque le groupe était soumis aux avis des autres, alors ils pouvaient s'allier plus facilement à un point de vue plutôt qu'à un autre.
- Cette étude permet aussi de réaliser que les professionnels ont besoin d'être formés sur les éléments qu'ils considèrent comme importants pour une exploitation éthique des données, mais donc leur compréhension s'écarte de celle de la littérature d'une part. On s'entend que si on n'a pas une bonne définition d'un élément, on n'aura pas une bonne application, car la base même n'est pas juste.
- En cas d'existence d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données, les gestionnaires gagneraient à expliquer aux scientifiques l'importance d'avoir une personne /un comité responsable et imputable de l'effectivité du code, car dans la perspective des scientifiques de données, une grande majorité est pour une gestion décentralisée ou chacun est responsable à son niveau. Certes, il y un avantage à ce que chacun soit responsable à son niveau, mais il y a un plus grand impact si cette action est conjointe à celle d'une personne imputable ayant pour mission de veiller à l'effectivité. Telle est d'ailleurs la position que les gestionnaires adoptent en grande majorité, à la suite de cette étude.
- Toujours au niveau pratique, cette étude permet d'identifier que le mode de contrôle interne et automatisé est favorisé par la majorité des experts

(scientifiques de données et gestionnaires). Par conséquent, une organisation qui mettrait sur pied un tel mécanisme pour contrôler l'effectivité des éléments importants à surveiller ne rencontrerait a priori pas de résistance de la part des professionnels.

- Enfin, bien que ce ne soit pas de façon naturelle, les professionnels sont ouverts à l'existence d'un mécanisme de contrôle externe qui travaillerait conjointement avec l'équipe interne pour l'effectivité d'un code balisant la pratique de leur métier au quotidien.

b) Contributions théoriques

Sur le plan théorique, on dénote 2 contributions essentielles de cette étude :

- Au niveau de la littérature, certains éléments considérés importants pour l'exploitation éthique des données ne sont pas suffisamment traités. Il s'agit entre autres des concepts : secret, exactitude, confidentialité où les professionnels (scientifiques de données et gestionnaires) ont des points de vue divergent en majorité, mais qu'au niveau de la littérature existante, il n'existe pas suffisamment de travaux qui traitent la question de manière à permettre à ceux-ci de s'y référer pour améliorer leur compréhension. À l'exemple de Foley et Guillemette qui ont mené une étude sur « qu'est-ce que l'intelligence d'affaires » en 2010, des travaux similaires pourraient être faits dans le but de bien définir ces termes, puis les vulgariser afin d'en assurer un bon usage par les professionnels de l'intelligence d'affaires et analytique. En effet, si la littérature estime qu'ils sont importants pour une exploitation éthique des données et si les professionnels s'accordent pour confirmer cela sur le plan pratique, il convient naturellement de s'assurer qu'ils sont bien définis, bien compris et utilisés adéquatement afin de garantir l'exploitation éthique des données dans toutes les dimensions considérées.
- Cette étude met sur pied les bases théoriques de la définition d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires dans un contexte d'analyse et d'exploitation de données vu sous l'angle des professionnels en intelligence d'affaires et analytique. Elle intègre à la fois les éléments importants à

considérer, les mécanismes de contrôle, le mode fonctionnement de ces mécanismes et les potentiels rôles liés à la responsabilité de l'effectivité du code.

4. Limites de la recherche

Les limites de cette recherche peuvent être perçues à deux niveaux : Méthodologie et analyse.

Au niveau de la méthodologie, on compte deux limites de cette étude. Premièrement, les constations se fondent sur les avis d'un nombre limité de professionnels et la taille n'est pas nécessairement représentative de la population des scientifiques de données ou des gestionnaires d'équipe d'intelligence d'affaires et analytique. Il n'est par conséquent pas prudent de généraliser les résultats de cette étude. Cependant étant donnée la nature exploratoire de la recherche, il importait de bâtir avec les professionnels les bases mêmes d'un code d'éthique et de déontologie qui baliseraient la pratique de leur métier. Deuxièmement, il a été mentionné dans la description de la méthode que celle-ci est longue et fastidieuse entraînant des abandons de la part des experts. C'est ce qui a été observé dans le cadre de cette étude. On a initialement contacté plus de 165 experts, mais finalement, 14 experts ont participé à l'enquête. On aurait pu avoir une troisième itération afin de préciser certains aspects liés à la gouvernance du code d'éthique dans un contexte d'exploitation des données, mais étant donné que les travaux sont effectués dans le cadre d'une maîtrise et non d'une étude de troisième cycle, on a fait le choix d'arrêter à deux itérations parce que la convergence était observée dans la majorité des éléments en étude.

Au niveau de l'analyse, la question 2 du premier tour de l'enquête demandait aux experts de proposer les autres éléments qu'ils trouvaient important d'intégrer dans un code d'éthique lié à l'exploitation des données. Plusieurs propositions ont été faites, mais n'ont pas été incluses dans le second tour premièrement parce qu'ils étaient différents les uns des autres et il y a en avaient autant qu'il y a de répondants. Deuxièmement, ces éléments n'étaient pas toujours liés aux enjeux éthiques en intelligence d'affaires et analytique dans une perspective d'exploitation de données, mais parfois relevaient plus de la gouvernance d'un code d'éthique et de déontologie

que de l'exploitation éthique des données. Troisièmement, sur la base des deux premières raisons, inclure ces éléments devait contribuer à alourdir l'enquête et potentiellement augmenter le taux d'abandon des répondants. Cependant, les différentes propositions sont disponibles à la section annexe (annexe 9 et annexe 12) du présent document.

5. Recherches futures

Dans cette section, il est question de proposer des ouvertures vers les recherches futures. À cet effet, quatre propositions sont retenues :

- On a proposé dans la section quelques conclusions importantes des recherches soient faites dans le but de définir certains concepts considérés importants pour l'exploitation éthique des données, mais pour lesquels les professionnels ont des points de vue différents quant à leur sens d'une part et d'autre part où la littérature n'est pas suffisamment fournie sur la question. C'est le cas d'exactitude, confidentialité et secret.
- Il a été annoncé dans la section problématique que la présente étude est de nature exploratoire et de ce fait, on a utilisé une méthodologie adaptée aux études de telle nature. Cependant, ces méthodes ont des inconvénients comme la taille des répondants qui pourrait être de nature à limiter l'application puisque l'avis de quelques experts vaut pour le groupe considéré. De ce fait, on pourrait sur la base de la présente étude, établir les principes de base de fonctionnement d'un code d'éthique et de déontologie dédiée à l'intelligence d'affaires et analytique dans un contexte d'exploitation des données puis utiliser la méthode par sondage pour valider auprès des professionnels en intelligence d'affaires et analytique ces principes. On pourrait ainsi attribuer un poids selon l'échelle de Likert à 7 valeurs pour chaque principe par exemple dans le but premièrement de valider les éléments qui ressortent de cette étude exploratoire.
- Étant donné que les organisations sont au balbutiement d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique comme mentionné pour le gouvernement fédéral du Canada, ou pour le cas des professionnels de

Montréal représentés par le groupe data science Montréal, où des initiatives sont faites, mais sans être des codes finis et adoptés en entreprise, une autre avenue de recherche pourrait être la mise en place d'un code d'éthique dédié à l'intelligence d'affaires dans un contexte d'analyse et d'exploitation de données, sur la base des résultats de la présente étude.

- Étant donné que la plupart des scientifiques de données n'a pas fourni de réponses lorsqu'on a demandé leur avis sur l'importance des termes pour une exploitation éthique des données, on pense qu'il serait intéressant de valider si l'éthique est enseignée dans leur programme, si oui comment, et ainsi voir si l'éthique est une avenue pour les scientifiques de données.

CHAPITRE 6 : CONCLUSION

La présente étude a été essentiellement motivée par l'absence d'un code d'éthique et de déontologie balisant la profession de l'intelligence d'affaires et analytique contrairement aux autres métiers du domaine des affaires. Ceci a abouti à la problématique de détermination des priorités à considérer lors de l'élaboration d'un code d'éthique dans le domaine de l'intelligence d'affaires et analytique afin de prémunir l'organisation face aux problèmes que pourrait engendrer une pratique non éthique de l'intelligence d'affaires et de l'analytique. Pour se faire, l'étude a été menée avec pour objectif de poser les bases conceptuelles pour l'élaboration d'un code d'éthique adapté, répondant aux impératifs des professionnels du domaine en l'absence de balises actuelles encadrant l'exploitation éthique des données malgré l'engouement de plus en plus croissant pour les données. À travers cet objectif principal, deux sous objectifs étaient de déterminer les principes fondateurs selon la perspective des professionnels et déterminer le mode de gestion à privilégier par ces mêmes professionnels. Pour atteindre cet objectif, la question de recherche de cette étude a été formulée ainsi qu'il suit : quels sont les enjeux que doit aborder un code d'éthique et de déontologie dédié à l'intelligence d'affaires et l'analytique selon les priorités perçues par les professionnels et les gestionnaires du domaine ? Pour arriver à cette fin, il était impératif de questionner les professionnels de l'intelligence d'affaires et analytique afin d'établir une priorité selon leur perspective. On a examiné le contexte d'étude et avons choisi la méthodologie qui convient le mieux à ce genre d'étude. Après analyse, la méthodologie DELPHI a été retenue du fait de la nature exploratoire de l'étude. Par la suite, conformément aux exigences de recherche avec les êtres humains, on a sollicité et obtenu l'accord du comité éthique de l'université de Sherbrooke. Une fois l'accord obtenu, il était possible de débiter effectivement en respectant toutes les étapes préconisées par la méthodologie adoptée. L'analyse des résultats de l'enquête menée a permis de valider auprès des professionnels les éléments importants liés à l'exploitation éthique des données qui sont soulignés au niveau de la littérature notamment, l'anonymat, la confidentialité, l'exactitude, le secret, l'intégrité, la transparence et la sécurité. Cette validation permet d'atteindre le premier sous objectif qui consiste à déterminer les principes fondateurs selon la perspective des professionnels. L'analyse a

aussi permis de déceler chez les professionnels l'habileté à se rallier lorsqu'ils sont exposés aux avis des autres. La poursuite de l'analyse permet de réaliser que les avis ne sont pas toujours convergents entre les deux groupes de professionnels du domaine de l'intelligence d'affaires et analytiques qui composent la population questionnée pour cette étude. Cette divergence permet de voir où les uns et les autres se situent au niveau du modèle de maturité de l'intelligence d'affaires et analytique selon Gartner (2008). L'analyse des résultats permet aussi de réaliser qu'au niveau des professionnels, le mode gestion de l'éthique et la déontologie ne fait pas l'unanimité. En effet, la population composée de gestionnaire d'équipe d'intelligence d'affaires et analytique est majoritairement en faveur d'un mode de gestion centralisé auprès d'un responsable qui veille à l'effectivité du code. Tandis que chez les scientifiques de données les avis sont en majorité partagés. On note également que les professionnels sont ouverts à des mécanismes de contrôle externe inspiré de celui des comptables professionnels agréés. Cette observation renseigne aussi sur le désir des professionnels d'adhérer à une pratique éthique de leur profession. Il en est de même de leur désir d'avoir un outil automatisé comme moyen de contrôle de l'éthique.

Au terme de cette étude, on peut dire que les objectifs escomptés sont atteints et ceci à l'intérieur des limites préalablement fixées tant au niveau méthodologie qu'au niveau analyse. En effet, un codage formel n'a pas été utilisé pour analyser les réponses comme c'est souvent le cas en analyse qualitative, aussi, certains éléments n'ont pas été intentionnellement analysés notamment les autres éléments que les professionnels considéraient importants pour une exploitation éthique des données. On peut cependant dire qu'il importe d'aller de l'avant et penser à la mise en place d'un code d'éthique et de déontologie dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données sur la base des résultats obtenus.

BIBLIOGRAPHIE:

Ackoff (1989); From data to wisdom. *Journal of Applied Systems Analysis* 1989; 16:3–9.

Adler M., Ziglio E. (1996); *Gazing into the oracle: The DELPHI method and its applications to social policy and public health*. Jessisca Kingsley Publishers Ltd 1996; p.252.

Alter, S., *Decision Support Systems* (1980); *Current Practices and Continuing Challenges*. Reading, MA: Addison-Wesley 1980.

Baillette, Paméla, Fallery, Bernard, et Girard, Aurélie (2013) ; *La méthode DELPHI pour définir les accords et les controverses : applications à l'innovation dans la traçabilité et dans le e-recrutement*. 18e Colloque de l'Association Information et Management (AIM) 2013 ; Lyon, France ; 1-22.

Baptendier, P. (2008); *Allez y on vous couvre! Un barbouze au service de l'État*. Editions du Panama 2008.

Bégin, Luc (2015) ; *L'éthique et les pratiques d'intervention en organisation*. Presses de l'Université Laval 2015.

Bergadaà, M. (2004) ; *Évolution de l'épistémè économique et sociale: Proposition d'un cadre de morale, de déontologie, d'éthique et de responsabilité pour le marketer*. *Recherche et Applications en Marketing* 2004; 19(1): 55-72.

Besson, B., Possin, J-C. (2001); *Du Renseignement à l'Intelligence économique*. 2e édition 2001; Dunod.

Bolger F., Wright G. (1994); *Assessing the quality of expert judgment, Decision support systems*, London. Jessisca Kingsley Publishers Ltd 1994; 11(1): 1-24.

Booto, Ekionea, Bernard et Plaisent (2011) ; *Consensus par la méthode DELPHI sur les concepts clés des capacités organisationnelles spécifiques de la gestion des connaissances*. *Recherches qualitatives* 2011; 29(3): 168-192.

Calvez, J. (2010); *Ethique, Morale, Déontologie : Proches et Différentes*. *Revue Française du Marketing* 2010; 226: 10-14.

Codd, E.F., Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate. Palo Alto, CA: E.F. Codd, 1993.

Crane, A. (2005); In the company of spies, when competitive intelligence gathering becomes industrial espionage. *Business Horizons* 2005; 48: 233-240.

Dalkey, N., & Helmer, O. (1963); An experimental application of the DELPHI method to the use of expert. *Management Science* 1963; 9(3): 458-467.

Davenport, Thomas H., & Harris, Jeanne G. (2007); *Competing on analytics*. Harvard Business School 2007; 46.

Delamotte, B. (2009); *Les coulisses de l'intelligence économique*. Nouveau Monde 2009.

Demilia, B., Peded, M., Jorgensen, K., & Subramanian, R. (2012). The ethics of BI with private and public entities. *Communications of the IIMA* 2012; 12(2):13-37.

Dhar, Vasant (2013); *Data Science and Prediction*. *Communications of the ACM* 2013 ; 12 : 64-73.

Dussureaul, David Gauthier (2015); *portrait de l'intelligence compétitive dans les petites et moyennes entreprises québécoises : comment faire de l'intelligence intelligente?* Université de Sherbrooke.

Ee-Peng, Lim, Hsinchun, Chen, & Guoqing, Chen (2012); *Business intelligence and analytics: Research directions*. *ACM Transactions on Management Information Systems* 2013; 17: 10 pages.

Evans, James R., Carl, H. (2012); *Business Analytics: The Next Frontier for Decision Sciences*. Lindner College of Business, University of Cincinnati, Decision Institute 2012.

Fayyad, U.M. et al., (1996); *Advances in Knowledge Discovery and Data Mining*. Cambridge, MA: MIT Press 1996.

Foley,Éric, Guillemette, Manon G. (2010); *What is Business Intelligence?* *International Journal of Business Intelligence Research* 2010; 1(4): 1-28.

Gartner Top 10 Strategic Technology Trends for 2019, David Cearley, Brian Burke, October 15, 2018.

Hopkins, Michael S., LaValle, Steve, Balboni, Fred, Kruschwitz, Nina, & Shockley, Rebecca. (2010); 10 data points: Information and analytics at work. MIT Sloan Management Review 2010; 52(1): 27-31.

Huff, Darrell (1954); How to Lie with Statistics. W. W. Norton & Company, 1954.

Ingham, John (2016) cours de méthodologie de recherche en intelligence d'affaires. Session 4.

Inmon, W.H. (1992); Building the Data Warehouse. New York, NY: Wiley, 1992.

Jordan, J., Finkelstein, and S. (2005); The ethic of competitive intelligence: Business case. Tuck School of Business at Dartmouth 2005.

Landeta, J. (2006); Current validity of the DELPHI method in social sciences. Technological forecasting and social change; 2006; 73(5):467-482.

Lanier, Paul, Rodriguez, Maria, Verbiest, Bryant, Sarah Katherine, Ting Guan, Zolotor, Adam (2019); Preventing Infant Maltreatment with Predictive Analytics: Applying Ethical Principles to Evidence-Based Child Welfare Policy. Journal of Family Violence 2019.

Lesparre, Josée (2019) ; Data analyst, Scientifiques des données. Sur le Web <https://www.cidj.com/metiers/data-analyst-data-scientist>, consulté le 12 avril 2019.

Liberatore, Matthew J., & Luo, Wenhong. (2010); The analytics movement: Implications for operations research. Interfaces 2010; 40(4): 313-324.

Linstone, H. A., Turrof, M. (1975); The DELPHI method, techniques and applications. Addison Wesley publishing 1975.

Linstone, Harold A. & Turoff, Murray (2002); The DELPHI Method, Techniques and applications. New Jersey Institute of Technology 2002.

McClave, James T. & co. (1997); Statistics. Prentice Hall 1997.

Moore, D. S. (1991); Teaching statistics as a respectable subject. In Gordon, F. and S. Gordon (eds.), *Statistics for the twenty-first century*, 14-25, Mathematical Association of America 1991.

N. K. McBride (2014); Active Ethics: an information systems ethics for the internet age. *Journal of Information Commun. Ethics Soc.* 2014; 12: 21–44.

Negash S., Gray P. (2008); Business Intelligence. In: *Handbook on Decision Support Systems 2. International Handbooks Information System*. Springer, Berlin, Heidelberg 2008, sur le Web : https://doi.org/10.1007/978-3-540-48716-6_9.

Okoli, C., & Pawlowski, S. D. (2004); The DELPHI method as a research tool: an example, design considerations and applications. *Information & Management* 2004; 42: 15-29.

Oueslati, Mohamed (2018); Facteurs d'influence de la maturité en intelligence d'affaires des entreprises dans les pays émergents : Le cas de la Tunisie. Université de Sherbrooke.

Paulsen & Coulson (2011); Beyond Awareness: Using Business Intelligence to Create a Culture of Information Security. *Communications of the IIMA* 2011; 11 (3).

Ricœur, P.(1990) ; Soi-même comme un autre. Le Seuil 1990.

Rojot, J. (2010); Déontologie et Gestion des Ressources Humaines. *Revue de Gestion des Ressources Humaines* 2010; (76): 31-43.

Rowe, G., Wright, G. (1999); The DELPHI technique as a forecasting tool: issues and analysis. *International Journal of Forecasting* 1999; 15: 353-375.

Saltz, Jeffrey S., Dewar, Neil (2019); Data science ethical considerations: a systematic literature review and proposed project framework. *Ethics and Information Technology* 2019; 21:197–208.

Stark, Luke and Hoffmann, Anna Lauren (2019); Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture. *Journal of Cultural Analytics* 2019.

Tan Ming Ming, Marzanah a. Jabar , Fatimah Sidi, Koh Tieng Wei (2015); A Systematic Literature Review Of Computer Ethics Issues. Journal of Theoretical and Applied Information Technology 2015; 78(3).

Thomann, J. B., & Wells, D. L. (2013); Managing Ethics in Business Intelligence. Business Intelligence Journal 2013; 18(2): 18-27.

Turban, E., Sharda, R., Aronson, J. E., and King, D. (2008); Business Intelligence: A Managerial Approach. Pearson Prentice Hall 2008.

Turilli, M. and Floridi, L. (2009); The ethics of information transparency. Ethics and Information Technology 2009; 11:105-122.

Vallejos, Elvira Perez, Koene, Ansgar, Carter, Christopher James, Hunt, Daniel Woodard, Christopher, Urquhart, Lachlan, Bergin, Aislinn, Statache, Ramona (2019); Accessing Online Data for Youth Mental Health Research: Meeting the Ethical Challenges. Philos. Technol. 2019; 32: 87.

Van Steenberg, Damien (2017), cours GIS 803 — Collecte de données, Usherbrooke Session d'été 2017.

Watson, H. J. (2012); This isn't your mother's BI architecture. Business Intelligence Journal 2012; 17(1): 4–6.

Watson, H. J. and Wixom, B. H. (2007); The current state of business intelligence. IEEE Computer 2007; 40(9): 96–99.

Wixom, B. H., Watson, H. J., and Werner, T. (2011); Developing an enterprise business intelligence capability. MIS Quartly 2011; 10(2).

Association Marketing Canadienne (ACM) 2019; sur le Web <https://www.the-cma.org/french/guides-de-conformite/code-de-deontologie-et-normes-de-pratique> consulté le 11 août 2019.

Association Marketing Canadienne (ACM) 2019; sur le Web <https://www.the-cma.org/regulatory/code-of-ethics> consulté le 11 août 2019.

Avant-premiere, 2019; sur le Web <https://avant-premiere.ca/ethique-et-intelligence-artificielle-ses-usages-au-canada/> consulté ce 14/05/2019 à 13h15.

Commission européenne. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), sur le Web https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC Consulté le 2018/08/27.

CPA 2019; sur le Web <https://www.cpacanada.ca/fr/connexion-et-nouvelles/blogue/articles-cpa/frise-chronologique-de-cpa-canada/frise-chronologique-de-cpa-canada> consulté le 11 août 2019

CRHA 2019 ; sur le Web <https://ordrecrha.org/ordre/ordre/historique> consulté le 10 août 2019.

CRHA 2019 ; sur le Web <https://ordrecrha.org/ordre/ordre/mission-valeurs-declaration-services> consulté le 10 août 2019.

DataDrivenMTL, 2019; sur le Web <https://www.meetup.com/fr-FR/DataDrivenMTL/> consulté le 14 mai 2019.

DSDTMTL, 2019 ; sur le Web <https://www.meetup.com/fr-FR/DSDTMTL/> consulté le 14 mai 2019.

FEPIE 2019; sur le Web <http://www.fepie.com/admin/docs/164.pdf> consulté le 10 août 2019.

Fobes, 2012 ; sur le Web <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#5dacaac06668> consulté le 14 août 2019.

Gouvernement fédéral du Canada 2019; sur le Web <https://www.canada.ca/fr/gouvernement/systeme/gouvernement->

numerique/technologiques-modernes-nouveaux/utilisation-responsable-ai.html consulté le 10 août 2019.

Radio-Canada 2018; sur le Web <https://ici.radio-canada.ca/nouvelle/1090381/scandales-donnees-facebook-cambridge-analytica-vises-3-enquetes-internationales> consulté le 14 août 2019.

Radio-Canada 2019 ; sur le Web <https://ici.radio-canada.ca/nouvelle/1193006/caisses-populaires-desjardins-vol-donnees-personnelles> consulté le 14 août 2019.

Radio-Canada 2019 ; sur le Web <https://ici.radio-canada.ca/nouvelle/1234371/facebook-amende-record-protection-donnees-personnelles> consulté le 08/08/2019.

SCIP 2019 ; sur le Web <https://www.scip.org/page/CodeofEthics> consulté le 10 août 2019

SCIP 2018 ; sur le Web <https://www.scip.org/page/CodeofEthics> consulté le 28.07.2018

SYNFIE, 2018; sur le Web <http://synfie.fr/synfie/gouvernance-synfie> consulté le 18 août 2018.

LEXIQUE :

IA&A : L'intelligence d'affaires et l'analytique

SQL: Structured Query Language

BI: Business Intelligence

IA : Intelligence artificielle

CAO : Chief Analytics Officer

CDAO : Chef officier de la donnée et de l'analytique

CDO : Chief Data Officer

CEST : Commission de l'éthique en science et en technologie

CNESST : Commission des normes, de l'équité, de la santé et de la sécurité du travail

CPA : Comptables Professionnels Agréés

DSS : Decision Support System

ETC : Extraction, transformation et chargement des données

ETL : Extract, Transform, Load

FEPIE : Fédération des Professionnels de l'Intelligence économique

IA : Intelligence d'affaires

IFPS : langages de planification financière de quatrième génération

LPRPDE : Loi sur la protection des renseignements personnels et les documents électroniques

OLAP : Online Analytical Processing (Traitement analytique en ligne)

PME : Petites et moyennes entreprises

RGPD : Règlement général sur la protection des données personnelles

RSE : responsabilité sociétale des entreprises

SYNFIE : Syndicat français de l'Intelligence économique

VP : Vice-Président

IAM : Identity Access Management

ANNEXES

Annexe 1 : lettre d'invitation

Samedi 15 décembre 2018

Bonjour M. X,

Je suis Chimène Flore FADJIE NGAYAP, étudiante au programme de maîtrise en administration concentration stratégie de l'intelligence d'affaires de l'Université de Sherbrooke.

Je sollicite votre participation, à titre d'expert, à une étude scientifique sur l'identification des contenus prioritaires d'un code d'éthique pour l'intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation en contexte de données massives selon la perspective des experts du domaine.

Votre participation consiste à donner votre avis d'expert sur la thématique de l'éthique en intelligence d'affaires et analytique au moyen d'un questionnaire échangé par courriel. Selon la méthode DELPHI, vous serez également appelé à vous prononcer sur les idées de l'ensemble des participants à quelques reprises afin d'établir un consensus sur les priorités des contenus d'un code d'éthique dédié à votre domaine d'expertise. L'enquête devrait durer approximativement 6 semaines. À l'issue du processus, vous recevrez un rapport synthèse des résultats.

Le résultat de cette recherche permettra d'identifier et de formaliser ces enjeux clés en intelligence d'affaires et analytique pour les diffuser auprès des chercheurs en académie et des praticiens du domaine afin d'améliorer l'encadrement des pratiques et prévenir les abus informationnels.

Contactez-moi par courriel (chimene.flore.fadjie.ngayap@usherbrooke.ca) afin de signifier votre intérêt pour l'étude. Je vous ferai alors parvenir les documents de projet. Par ailleurs, n'hésitez pas à me contacter si vous avez des questions sur le projet.

Votre acceptation à ce projet aidera notre étude et l'avenir du champ de recherche.

Je vous en remercie,

Chimène Flore FADJIE Ngayap

Annexe 2 : Courriel pour l'invitation

Objet : Participation à une étude sur l'éthique en BI

Contenue :

Bonjour M.

Vous trouverez ci-joint les documents relatifs à l'étude portant l'éthique en Intelligence d'affaires.

Il s'agit d'une :

- **Lettre d'invitation à participer à l'enquête** qui indique en quoi consiste votre participation ainsi que l'importance de votre participation;
- **Lettre de consentement qui vous donne plus de détails sur le projet de recherche.** Je vous prie d'en prendre connaissance, y apposer votre signature et nous le retourner avec votre réponse au questionnaire ;
- **Le questionnaire pour l'enquête.**

Le questionnaire vous a été soumis sous format Word afin de permettre plus de flexibilité dans la mesure où vous pourrez répondre aux questions à des moments différents et nous retourner le tout seulement quand vous aurez terminé.

Je vous remercie.

Chimène Flore Ngayap

Annexe 3 : message de sollicitation via linkedIn

Bonjour,

Professionnelle en BI et étudiante à Usherbrooke, j'effectue une étude dirigée par le Dr Daniel Tremblay sur l'éthique en BI et je sollicite la participation des professionnels comme vous. Serez vous intéressé à y participer? Si oui, je vous enverrai plus de détails par courriel.

Merci.

Annexe 4 : Formulaire d'information et de consentement

Vous êtes invité(e), à titre d'expert, à participer à un projet de recherche universitaire en intelligence d'affaires et analytique. Le présent document vous renseigne sur les modalités de ce projet de recherche. S'il y a des mots ou des paragraphes que vous ne comprenez pas, n'hésitez pas à contacter la responsable du projet pour poser vos questions. Pour participer à ce projet de recherche, vous devrez apposer votre signature à la dernière page de ce document et nous le faire parvenir. Nous vous remettrons une copie signée et datée du document et conserverons l'original pour la durée du projet. Prenez tout le temps nécessaire pour prendre votre décision.

Projet

Titre du projet de recherche

Perspective des professionnels de la donnée sur les contenus prioritaires d'un code d'éthique pour l'intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation en contexte de données massives.

Remarque :

1. *Nous définissons le code d'éthique dédié à l'intelligence d'affaires et l'analytique comme étant un texte énonçant les valeurs et les principes à connotation morale ou civique auxquels adhère une organisation et qui servent de guide aux professionnels de l'intelligence d'affaires et de l'analytique notamment les scientifiques de données, les intendants de données, les analystes de données des lignes d'affaires afin de les aider à juger de la justesse de leurs comportements.*
2. *La définition ci-dessus évoquée s'applique non seulement à ce document, mais aussi à tous les documents où apparaîtra le terme code d'éthique dans le contexte de notre projet de recherche.*

Personnes responsables du projet de recherche

La responsable du projet est Chimène Flore FADJIE NGAYAP étudiante en maîtrise en administration des affaires à l'Université de Sherbrooke. Vous pouvez la rejoindre en composant le (xxx) xxx - xxxx, ou par courriel à chimene.flore.fadjie.ngayap@usherbrooke.ca. Son travail est effectué sous la direction de M. Daniel CHAMBERLAND-TREMBLAY.

M. Daniel CHAMBERLAND-TREMBLAY est professeur agrégé à l'École de gestion de l'Université de Sherbrooke. On peut le joindre au (819) 821-8000 poste 61316 (ou sans frais au 1 800 267-8337) ou par courriel à daniel.chamberland-tremblay@usherbrooke.ca.

Objectifs du projet de recherche

L'objectif principal de ce projet est d'identifier les fondements du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives *selon la perspective des professionnels du domaine*.

La recherche permettra de définir, selon la perspective des professionnels du domaine, les enjeux prioritaires à considérer et à traiter dans le contexte émergent des données

massives où le volume et la diversité des données et les capacités de traitements informatiques soulèvent des problèmes comme la confidentialité et l'anonymisation des données et pour lequel il n'existe pas actuellement de modèle dédié.

Raison et nature de la participation

À titre d'analyste des données dans une ligne d'affaires, de scientifique des données ou d'expert en analytique, vous êtes invité(e) à partager votre avis de spécialiste sur les contenus fondateurs d'un code d'éthique dédiée à l'intelligence d'affaires et l'analytique pour soutenir les pratiques d'analyse et d'exploitation des données massives dans les organisations.

Ce projet s'appuie sur une démarche formelle d'étude, nommée enquête DELPHI, qui sollicite votre avis d'expert pour d'abord identifier les contenus d'un code d'éthique dédiés à votre contexte d'affaires, puis de les prioriser.

À cet égard, votre participation à ce projet consiste à répondre à un questionnaire initial qui vous sera transmis par courriel. Répondre à ce questionnaire demandera moins de 60 minutes.

Par la suite, vous devrez vous prononcer sur l'avis de l'ensemble des experts ayant participé à l'enquête, et ce, dans le but de faire émerger un consensus sur les contenus prioritaires pour votre domaine d'expertise. Cette étape repose sur un questionnaire modifié qui vous sera également envoyé par courriel et qui ne prendra que quelques minutes à répondre. Vous pourriez être sollicité(e) à quelques reprises (deux itérations sont actuellement prévues) afin d'assurer la convergence des avis d'experts vers un consensus.

L'ensemble de la recherche se déroulera sur une période d'environ 6 semaines.

Avantages pouvant découler de la participation

Votre participation à ce projet de recherche vous permettra d'obtenir en exclusivité une synthèse des résultats de la consultation auprès des experts de l'intelligence d'affaires et de l'analytique.

Aussi, votre participation contribuera à établir une compréhension formelle plus large des contenus prioritaires d'un code d'éthique pour l'intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation en contexte de données massives dans une perspective d'avancement de la science.

Inconvénients et risques pouvant découler de la participation

Votre participation à cette étude nécessite l'investissement de votre temps professionnel ou personnel. Toutefois, l'étude est sans risque pour les répondants.

Participation volontaire et possibilité de retrait

Votre participation à ce projet de recherche est volontaire. Vous êtes donc libre de refuser d'y participer. Vous pouvez également vous retirer de ce projet à n'importe quel moment, sans avoir à donner de raisons, mais simplement en informant la responsable du projet de recherche. Advenant votre retrait de l'étude, l'information déjà recueillie sera conservée, analysée et utilisée de façon anonyme pour assurer l'intégrité du projet de recherche.

Confidentialité, partage, surveillance et publications

La chercheure vous a avisé que cette étude ne comporte aucun risque, puisque toute l'information sera traitée de façon confidentielle. Par exemple, dans le rapport relatif à chaque tour de la consultation DELPHI, les références à vos réponses seront anonymisées et présentées sous la forme « Expert 000 », plutôt que de préciser votre nom et votre fonction.

Durant votre participation à ce projet de recherche, la chercheure responsable recueillera, dans un dossier de recherche, les renseignements vous concernant et nécessaires pour répondre aux objectifs scientifiques de ce projet de recherche.

Votre dossier de recherche peut comprendre des renseignements tels que votre nom, votre sexe, votre date de naissance, votre employeur et vos années d'expérience en qualité d'analyste des données dans une ligne d'affaires, de scientifique des données ou d'expert en analytique.

Tous les renseignements recueillis au cours du projet de recherche demeureront strictement confidentiels dans les limites prévues par la loi et vous ne serez identifié(e) que par un numéro de code. La clé du code reliant votre nom à votre dossier de recherche sera conservée par la chercheure responsable de ce projet de recherche.

Les données recueillies dans le cadre de ce projet de maîtrise seront conservées, sous clé, pendant 5 ans suivant la date de dépôt du mémoire par la chercheure responsable aux fins exclusives du présent projet de recherche puis détruites.

Les données de recherche, dans leur forme anonymisée, pourront être publiées ou faire l'objet de discussions scientifiques, mais il ne sera pas possible de vous identifier.

À des fins de surveillance et de contrôle, votre dossier de recherche pourrait être consulté par une personne mandatée par des organismes réglementaires, des représentants de l'établissement ou du comité d'éthique de la recherche. Ces personnes et ces organismes adhèrent à une politique de confidentialité.

Vous avez le droit de consulter en tout temps les données recueillies pour ce projet de recherche pour vérifier les renseignements recueillis et les faire rectifier au besoin.

Résultats de la recherche

Comme participant à l'étude, vous serez informé(e) en priorité, avant toute diffusion scientifique, des résultats de la recherche et des publications qui en découleront le cas échéant.

Coordonnées de personnes-ressources

Si vous avez des questions ou si vous éprouvez des problèmes reliés au projet de recherche, ou si vous souhaitez vous en retirer, vous pouvez communiquer avec la chercheure responsable au numéro suivant : (xxx) xxx xxxx.

Approbation par le comité d'éthique de la recherche

Le Comité d'éthique de la recherche - Lettres et sciences humaines de l'Université de Sherbrooke a approuvé ce projet de recherche et en assurera le suivi. Pour toute question concernant vos droits en tant que participant à ce projet de recherche ou si vous avez des commentaires à formuler, vous pouvez communiquer avec ce comité au numéro de téléphone 819-821-8000 poste 62644 (ou sans frais au 1 800 267-8337) ou à l'adresse courriel cer_lsh@USherbrooke.ca.

Signature de la personne participante

J'ai pris connaissance du formulaire d'information et de consentement. On m'a expliqué le projet de recherche et le présent formulaire d'information et de consentement. On a répondu à mes questions et on m'a laissé le temps voulu pour prendre une décision. Après réflexion, je consens à participer à ce projet de recherche aux conditions qui y sont énoncées.

Nom de la personne participante

Signature

Date

Engagement de la chercheure responsable du projet de recherche

Je certifie que j'ai expliqué à la personne participante le présent formulaire d'information et de consentement, que j'ai répondu aux questions qu'elle avait.

Je m'engage, avec l'équipe de recherche, à respecter ce qui a été convenu au formulaire d'information et de consentement et à en remettre une copie signée et datée à la personne participante.

Nom de la chercheure responsable

Signature

Date

Annexe 5 : Questionnaire 1^{er} tour pour l'enquête DELPHI

Objectifs du projet de recherche

L'objectif principal de ce projet est d'identifier les fondements du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives *selon la perspective des professionnels du domaine*.

Question 1 : Une revue de la littérature scientifique et professionnelle a permis d'identifier des concepts liés à l'utilisation éthique des données en intelligence d'affaires et analytique. Les principaux concepts répertoriés incluent l'anonymat, la confidentialité, la sécurité, le secret, l'exactitude, l'intégrité et la transparence. Considérez-vous ces éléments comme déterminant à la définition d'un code d'éthique spécifique à l'intelligence d'affaires dans un contexte d'exploitation de données ?

Pour chaque terme, nous vous prions d'indiquer le sens que vous attachez au terme et une brève description de son importance dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives.

Question 2 : Selon votre expertise, quels sont les autres concepts ou éléments qui doivent être pris en compte de manière prioritaires dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives ?

Pour chaque terme, nous vous prions d'indiquer le sens que vous attachez au terme et une brève description de son importance dans l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives.

Question 3 : En vous basant sur votre expertise, placez les éléments retenus à la Question 1 et à la Question 2 par ordre de priorité du plus prioritaire au moins prioritaire pour l'élaboration du contenu d'un code d'éthique en intelligence d'affaires et l'analytique dédié aux pratiques d'analyse et d'exploitation des données massives.

Question 4 : À titre d'expert de l'exploitation de données, qui dans l'organisation devrait être responsable de veiller à l'effectivité, c.-à-d. la mise en opération effective, mesurée et pérenne, de chacun des éléments retenus à la Question 1 et à la Question 2 ?

Question 5 : À titre d'expert de l'exploitation de données, quels sont les mécanismes de contrôle de gestion à considérer pour chacun des éléments retenus à la Question 1 et à la Question 2 ?

Annexe 6 : courriel accompagnant le 2^e questionnaire

Objet : Étude sur l'éthique en BI deuxième tour d'enquête

Bonjour,

Nous sommes ravis de vous écrire une fois de plus dans le cadre du projet de recherche sur l'éthique en intelligence d'affaires.

Nous vous remercions pour votre participation au premier tour du questionnaire. Plusieurs mois se sont écoulés depuis que celui-ci vous a été soumis ; ce délai a permis aux uns et aux autres de compléter et retourner leurs réponses, mais aussi de procéder à une première analyse.

Suite à cette analyse préliminaire, nous désirons vous entendre sur certains points qui restent à clarifier ; raison pour laquelle, nous vous sollicitons à nouveau.

Vous trouverez donc ci-joint deux documents :

- Une lettre de remerciement du directeur de cette recherche : le professeur Daniel Chamberland-Tremblay
- Le questionnaire pour le deuxième tour de l'enquête.

Comme la dernière fois, le formulaire vous est soumis en format Word afin de permettre une flexibilité pour la réponse. Ainsi, vous pourrez répondre aux questions à votre rythme et nous retourner le tout lorsque vous aurez répondu à toutes les questions.

Cordiales salutations.

Chimène Flore Ngayap

Annexe 7 : Lettre de remerciement du directeur d'étude



Sherbrooke, le 2 avril 2019

Madame/Monsieur,

C'est avec grand enthousiasme que je communique avec vous dans le cadre de cette étude exploratoire sur la définition et l'application des pratiques éthiques et déontologiques dans l'analyse de données d'entreprise à l'Université de Sherbrooke.

Tout d'abord, je souhaite vous remercier personnellement pour votre contribution à l'avancement de la recherche. Votre avis d'expert est précieux dans un contexte où s'ouvre un nouveau champ d'expertise dans un domaine en pleine évolution. Mes étudiantes Chénène Fiore Ngayap, responsable du projet, témoigne de votre situation à la lumière des résultats préliminaires :

- Bonnamment, il y a un nombre d'experts en Data Science qui ne se sentent que peu ou pas concernés par l'établissement de balises dans la pratique éthique ou déontologique de l'analytique. Bien que ce constat soit anecdotique et en marge de la présente étude, il réitère l'importance de votre contribution dans ce projet.
- Par ailleurs, on constate des points de vue et des compréhensions divergents par les experts des concepts fondamentaux de l'éthique en Data Science. Ce résultat intermédiaire était attendu puisque peu de formation en intelligence d'affaires ou analytique intègre les notions d'éthique d'utilisation des données. L'approche de recherche Delphi a été choisie spécifiquement pour permettre d'aborder cette situation. Le deuxième tour de consultation devrait permettre d'amener un éclairage intéressant sur les priorités perçues par les experts.
- Enfin, la gestion de l'éthique des données en Data Science est perçue selon deux approches distinctes qui, dans d'autres domaines connexes, sont souvent mises en opposition. Votre apport sur le sujet est crucial pour comprendre ce qui motive le choix d'un mode de gouvernance ou un autre.

Les données récoltées lors de la consultation initiale ont permis d'établir des constat préliminaires particulièrement riches et intéressants. Toutefois, il reste des zones d'ombre à éclaircir avant de pouvoir finaliser cette recherche préliminaire.

Je vous enjoins donc à compléter cette consultation complémentaire directement issue des constats préliminaires. Nous avons construit le document afin de maximiser votre contribution et de limiter au maximum le temps à y consacrer.

Je vous remercie d'avance de votre contribution.

Recevez mes salutations distinguées,

Daniel Chamberland-Tremblay
Professeur
Université de Sherbrooke

Annexe 8 : Analyse sommaire et questionnaire 2^e tour

Bonjour,

Nous souhaitons tout d'abord vous remercier d'avoir participé à l'établissement des fondements pour l'étude sur le rôle de *scientifique de données* dans la définition et l'application des pratiques éthiques et déontologiques dans l'analyse de données d'entreprise par l'Université de Sherbrooke. Votre avis, comme celui de tous les experts ayant participé, est précieux et éclairant.

En raison de la diversité des résultats reçus, nous vous contactons à ce stade critique de la recherche afin d'établir de façon concertée les priorités éthiques et déontologiques pour la Science des données et de son implantation en entreprise. Nous avons divisé ce document en sections pour faciliter la lecture, vous n'aurez besoin que de quelques minutes pour y répondre.

Constat 1 : Disparité dans la compréhension des termes liés à l'éthique

À l'exception du terme anonymat pour lequel l'ensemble des répondants a offert une réponse similaire, tous les autres termes présentent une disparité dans leur compréhension. Nous avons repris dans le tableau suivant les différentes réponses et nous vous invitons à mettre un 'X' dans la case qui correspond le mieux à votre compréhension du terme ou bien à fournir une nouvelle proposition sous la rubrique « Autre définition »

Termes	Définitions suggérées	Votre choix de réponse
Confidentialité :	▪ non divulgation de la donnée	
	▪ nonaccès ou restriction d'accès uniquement aux personnes autorisées	
	▪ garder secret l'information détenue et éviter que celle-ci permette d'identifier la personne référencée.	
Autre définition :		
Sécurité :	▪ pas d'accès environnemental aux données par des personnes non autorisées.	
	▪ sécuriser les données en préservant son intégrité ainsi que tout ce qui pourrait mettre la donnée en danger.	
Autre définition :		
Secret :	▪ combinaison de confidentialité, anonymat et sécurité.	

	<ul style="list-style-type: none"> ▪ ne pas divulguer ou dévoiler une information ou le but de l'analyse d'une donnée 	
	<ul style="list-style-type: none"> ▪ le sens n'est pas clair 	
Autre définition :		
Exactitude :	<ul style="list-style-type: none"> ▪ vérité, ou ce qui est conforme à la réalité. 	
	<ul style="list-style-type: none"> ▪ « exact » et démunie de toute falsification. 	
	<ul style="list-style-type: none"> ▪ « bonne » et utilisable. 	
	<ul style="list-style-type: none"> ▪ bonne qualité et la possibilité de s'y fier. 	
	<ul style="list-style-type: none"> ▪ précision et certitude de l'information dérivée de la donnée 	
Autre définition :		
Intégrité :	<ul style="list-style-type: none"> ▪ bonne intention de ceux qui traitent la donnée 	
	<ul style="list-style-type: none"> ▪ Le fait que la donnée garde son état initial durant le processus de traitement et ne subisse aucune altération. 	
	<ul style="list-style-type: none"> ▪ le fait d'intégrer la donnée dans un ensemble plus global avec possibilité que celle-ci soit validée en recoupant avec d'autres données références 	
Autre définition :		
Transparence :	<ul style="list-style-type: none"> ▪ énoncer clairement l'utilité finale de la donnée ainsi que la façon dont elles sont utilisées. 	
	<ul style="list-style-type: none"> ▪ rendre publique la provenance de la donnée. 	
	<ul style="list-style-type: none"> ▪ rendre accessible la donnée elle-même et non la source lorsqu'elle concerne l'opinion publique. 	
	<ul style="list-style-type: none"> ▪ documenter de façon complète dans un langage non codé les processus relatifs à la protection, exploitation et diffusion de l'information. 	
	<ul style="list-style-type: none"> ▪ transparence au niveau de l'analyse faite avec les données 	
Autre définition :		

Constat 2 : Éléments prioritaires dans la définition et la mise en œuvre d'un code d'éthique

Quels aspects doivent être priorités dans l'établissement d'un code d'éthique pour la science des données ? Dans le tableau suivant, vous êtes invité à indiquer selon un ordre de priorité compris entre 1 et 7 où 1 représente l'élément le plus prioritaire et 7 l'élément le moins prioritaire.

Termes	Ordre de Priorité compris entre 1 et 7 (où 1 est le plus prioritaire et 7 le moins prioritaire. Bien vouloir ne pas attribuer la même valeur à deux termes)
Anonymat	
Confidentialité	
Exactitude	
Intégrité	
Secret	
Sécurité	
Transparence	

Vous êtes invité à expliquer votre proposition, si vous le souhaitez.

--

Constat 3 : Gouvernance de l'éthique

Les résultats reçus présentent un clivage important sur la responsabilité de l'effectivité d'un code d'éthique dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données. Il nous importe de comprendre votre position à cet égard.

D'un côté, un groupe de répondants associe cette responsabilité au *Chief Data Officer* (DCO), au Chief Analytics Officer (CAO), ou au chef de la donnée et de l'analytique (CDAO) et propose de créer ce poste si l'entreprise n'en dispose pas ; ou encore l'associe à un haut cadre comme le VP marketing de l'entreprise.

De l'autre côté, un groupe équivalant de répondants impute la responsabilité de l'effectivité au gestionnaire d'un comité, ou d'une équipe, dédié à l'éthique en matière d'exploitation des données dans l'entreprise.

Les approches adoptent deux postures différentes de la gouvernance en intelligence d'affaires et analytique : centralisation et décentralisation. Selon vous, laquelle de ces deux propositions

serait la plus appropriée spécifiquement pour l'effectivité d'un code d'éthique dédié à l'intelligence d'affaires et analytique ?

Vous êtes invité à expliquer votre position sur ce sujet spécifique.

Constat 4 : mécanisme de contrôle

Afin de permettre la viabilité d'un code dédié à l'intelligence d'affaires et analytique dans un contexte d'exploitation de données, vous avez été appelé à vous prononcer sur les mécanismes de contrôle en entreprise. Deux propositions ressortent de l'analyse préliminaire des réponses de l'enquête.

La première proposition est un mécanisme paragouvernemental dédié : **une commission des normes éthiques en matière d'analyse et d'exploitation de données** responsable de s'assurer que des mesures et des dispositions relatives au respect d'un code éthique sont opérationnelles dans chaque entreprise effectuant de l'analyse de données de quelque forme que ce soit. Ce mécanisme s'apparente à la Commission des normes, de l'équité, de la santé et de la sécurité au travail (CNESST) qui effectue des contrôles périodiques, puis établit des responsabilités et des mesures correctives lorsqu'un écart est observé.

La deuxième proposition consiste à définir des procédés de contrôle interne à l'entreprise pour chaque aspect du code d'éthique : **mettre en place un mécanisme dynamique d'évaluation**, incluant les techniques d'apprentissage automatique, **pour la détection en continu des anomalies ou du non-respect des éléments clés d'un code d'éthique** appliqué à l'intelligence d'affaires et analytique

Les propositions mettent en opposition un contrôle interne et externe des pratiques des *Data Scientists*. Ces deux stratégies ne sont pas nécessairement exclusives, mais sont issues de préoccupations distinctes.

Afin de mieux comprendre votre réflexion à titre d'expert, nous aimerions savoir si vous partagez l'idée de la création d'une commission nationale pour l'éthique de l'exploitation des données en intelligence d'affaires et analytique.

Comment envisageriez-vous son fonctionnement ?

Seriez-vous favorable à l'élaboration de mécanismes de contrôle internes et automatisés ?

Comment envisageriez-vous son fonctionnement ?

D'après vous, les deux mécanismes de contrôle identifiés doivent-ils cohabiter ou est-ce redondant ? Vous êtes invité à expliquer votre position sur ce sujet spécifique.

Annexe 9 : Autres éléments considérés importants par les scientifiques de données et liés à l'exploitation éthique des données

autres termes à considérer proposer par les scientifiques des données		
Consentement	La non-discrimination	Le conflit d'intérêt:
Protection des données personnelles	La validation croisée	Le respect de la loi:
Conservation	Utiliser pour le besoin seulement	Le processus de dénonciation:
Doit de retrait	Procédés de Remédiation	Formation
Droit d'utilisation	Période de rétention	Évaluation
Accès	Adhésion aux Règlementations	Valeur
Documentation	Le consentement	La responsabilité:
Gouvernance		

Annexe 10 : Mécanismes de contrôle proposés par les scientifiques de données

<ul style="list-style-type: none"> • Anonymat, Confidentialité : L'équipe de gouvernance doit mettre en place un processus prenant place dès lors de l'intégration de nouvelles données à des fins analytiques. Ce processus doit déterminer le type de transformation(s) devant être appliquées aux données pour assurer l'anonymat et la confidentialité. Ces transformations devront être appliquées sur les données avant de les rendre disponibles plus globalement dans les entrepôts de données. • Sécurité, Secret: L'équipe de sécurité corporative doit s'assurer que tout groupe ayant accès à des données sensibles respecte les mesures nécessaires pour assurer la sécurité et le secret des données. Plus particulièrement, des règles doivent être mises en place pour assurer que les pratiques et logiciels informatiques respectent les standards de sécurités. • Formation : Des formations doivent être offertes après tout changement significatif dans les règles de confidentialité ou sécurité, ou bien lors de l'intégration de nouvelles pratiques ou nouveaux outils technologiques. Des formations récurrentes (annuelles) doivent être offertes pour permettre de faire un rappel. • Évaluation : Des contrôles 	<p>Mécanismes de contrôle de gestion à considérer</p> <ul style="list-style-type: none"> · Gestion des accès aux données. · Chiffrement des données. · Registre consignait les activités de traitement des données afin d'établir la traçabilité des traitements de données. · Audits de conformités sur les accès aux données. · Audits de conformités sur les traitements des données. · Audits de conformités sur le respect des lois, réglementations et politiques applicables aux traitements des données. 	<ul style="list-style-type: none"> • Anonymat <ul style="list-style-type: none"> o Gouvernance de données qui dicte la portée (quelles données) et les méthodes d'anonymisation des données (comment). Le contrôle qualité du traitement de données devrait avoir sur sa liste de validation un point sur l'anonymat et des audits devraient être performés pour s'assurer du respect de la gouvernance. • Confidentialité <ul style="list-style-type: none"> o Avoir une politique claire de comment les données recueillies seront traitées et s'assurer que cette politique est respectée. Un responsable des données devrait être identifié et il devrait être rigoureux quant à la gestion des données confidentielles dont il est responsable. La documentation des données devrait indiquer clairement si une donnée est confidentielle. L'utilisation possible ainsi que les rôles et responsabilités devraient être clairs pour ces données. • Sécurité <ul style="list-style-type: none"> o S'assurer d'avoir une politique de sécurité des données claire et la respecter. Des audits de sécurité peuvent être réalisés afin de s'assurer de la conformité à la politique. • Secret <ul style="list-style-type: none"> o Voir confidentialité (synonyme moi) • Exactitude <ul style="list-style-type: none"> o La gouvernance de données devrait avoir des règles claires à cet effet ainsi qu'un responsable de la donnée qui s'assure de l'exactitude et de l'intégrité des données dont il est responsable.
---	--	--

<p>ponctuels doivent être effectués.</p> <ul style="list-style-type: none"> • Exactitude, Intégrité : L'équipe de gouvernance doit mettre en place des systèmes d'évaluation de l'exactitude et de l'intégrité des données. À la fois à la source, mais aussi durant le cycle de vie complet des données. • Transparence : 		<p>Le contrôle qualité devrait s'assurer que les données ne sont pas dénaturées lors de leur traitement. La traçabilité des données « data lineage » permet également de rendre visible l'ensemble des traitements opérés sur une donnée ainsi que de remonter jusqu'à la source pour s'assurer qu'une donnée n'a pas été dénaturée.</p> <ul style="list-style-type: none"> • Intégrité <ul style="list-style-type: none"> o Voir exactitude (synonyme selon moi) • Transparence <ul style="list-style-type: none"> o Rendre disponible les traitements et la donnée source de chaque donnée utilisée. Ceci peut être fait à l'aide de la traçabilité des données « data lineage ».
--	--	---

<p>Mécanisme de contrôles:</p> <ul style="list-style-type: none"> - Avoir un système de traitement de plaintes géré par la commission des normes éthiques en matière d'analyse et d'exploitation de données. - Chaque organisation serait soumise à la possibilité constante d'être auditée par un agent certifié de la (théorique) commission des normes éthiques en matière d'analyse et d'exploitation de données dans le cas d'une ou plusieurs plaintes sérieuses. - Chaque organisation devrait se doter d'au moins une personne responsable de veiller à l'éthique en matière d'analyse et d'exploitation de données. 	<p>Anonymat : Ne demander et stocker que les informations strictement nécessaires sur les individus.</p> <p>Confidentialité : Avoir un document clair sur le partage des données sensibles. Intégrer ces notions dans la culture de l'entreprise.</p> <p>Sécurité : VPN, Formations annuelles sur les dangers et la prévention, pare-feu</p> <p>Secret : Avoir un document clair sur le partage d'informations sur la détention de données. Intégrer ces notions dans la culture de l'entreprise.</p> <p>Exactitude : QA (Quality Assurance) de ses analyses par soi-même et par ses pairs. Pouvoir QA certaines informations à partir d'autres documents.</p> <p>Intégrité : Éviter/Corriger toute répétition de données inutile (Data Manager). Garder les Data Manager au courant des données erronées trouvées (Analystes).</p> <p>Transparence : Toujours indiquer de façon claire la méthodologie effectuée pour en arriver à cette fin.</p>	<p>Mécanismes d'évaluations d'impacts (privacy impact assessment) et vérifications, pour adhérer aux éléments Défini par le 'Data Controller'.</p> <ul style="list-style-type: none"> • Anonymat • Confidentialité • Secret • Exactitude • Intégrité • Transparence • Période de rétention <p>Adhésion aux réglementations gouvernementales:</p> <ul style="list-style-type: none"> • Utiliser pour le besoin seulement • Procédés de remédiation • Procédés de corrections
	<p>Évolution de la qualité des données dans le temps</p> <p>L'usage des solutions BI</p> <p>Évolution de la culture analytique au sein de l'organisation</p> <p>Formation en BI des ressources</p> <p>Le BI comme une compétence distinctive</p>	

<p>Le mécanisme clé dans le contrôle de gestion en analytique est la transparence.</p>	<p>Sécurité : Propriétaire de données; Processus d'approbation Confidentialité « de l'information »: Formation; Evaluation; Signature Transparence : Processus de reddition; registre des méthodes; registre des sujets; registre des règles Secret : Information; organisation de tests d'évaluation; Évaluation réelle Gouvernance: Cadre de référence; Politique et normes; Contrôle périodique et aléatoire; Inscription du respect du code d'éthique dans la performance des individus, des départements d'une entité Intégrité : Normes professionnelles Exactitude: Normes professionnelles Anonymat : Non retenu pour ma part</p>	<p>Comme il y'a plusieurs concepts à surveiller avec un grand volume de données, d'accès et d'utilisation, je vois deux mécanismes : un qui est statique et l'autre est dynamique. Le mécanisme statique est l'audit, cette équipe doit s'assurer pour un projet/produit ou une application donnée que tous ces concepts sont respectés. Le mécanisme dynamique est l'utilisation de la sécurité analytique qui inclut les techniques d'apprentissage automatique pour la détection en continu des anomalies ou le non-respect de ces concepts.</p>
--	---	---

Annexe 11 : Avis des scientifiques de données au sujet de mode de gouvernance de l'éthique et la déontologie

Mode de gouvernance de l'éthique		
<p>Centralisation. Explication : Je pense que la centralisation permettrait d'assurer que les règles soient comprises de façon uniforme et mises en place de la même façon lors de l'usage des données peu importe la donnée, la source et leur traitement.</p>	<p>La centralisation. Explication: L'avantage de la centralisation avec un CDO comme sponsor permet de mieux mobiliser les différentes ressources requises à bonne marche au niveau corporatif et d'aligner les orientations avec celles de la stratégie corporative</p>	<p>Centralisation (CDO) Explication: L'important est de garder « one version of the truth ». On ne veut pas avoir plusieurs règles différentes pour gouverner différentes sphères de l'entreprise. Cela peut fonctionner pour un certain moment, mais éventuellement, on veut amener l'entreprise à adopter un modèle BI plus mature, ce qui veut également dire centraliser les efforts d'un même code d'éthique.</p>
<p>Je ne suis pas en mesure de trancher pour une ou l'autre. Explication: Je crois que cela dépend de trop de facteurs :</p> <ul style="list-style-type: none"> • L'organisation et sa culture • Le niveau de maturité • Le secteur d'activité (orienté données ou pas) • Etc. <p>Je crois simplement qu'une instance soit nécessaire pour s'assurer de l'effectivité mais celle-ci peut prendre diverses formes.</p>	<p>Une instance de surveillance externe à celles et ceux qui interviennent en intelligence d'affaires et analytique. Une instance Centrale ou fédérée selon les organisations, comme un comité d'éthique et juridique. Explication: Les réalités organisationnelles sont complexes (société mère vs filiales, organisation matricielle, activités diversifiées comme les conglomérats). La réponse devrait être fonction du contexte organisationnel, car l'existence des centres d'expertises sont à l'image des entités où ils existent et doivent aborder les aspects éthiques et juridiques selon les sources des données, les usages, et les législations applicables. Nul ne peut être juge de soit même et en matière d'éthique il ne peut y avoir d'autorégulation.</p>	<p>Je trouve que les deux propositions se complètent. Explication: La responsabilité et l'imputabilité revient à l'équipe de gouvernance de données avec leur lead DCO. Par contre, l'équipe de conformité ou d'audit s'assure que la gouvernance est bien respectée dans l'entreprise par l'organisation d'audit et du suivi de l'utilisation des données personnelles.</p>

<p>Je pense qu'il importe de jouer des deux approches.</p> <p>Explication: Savoir centraliser la stratégie reliée à l'éthique auprès d'un CDO et savoir décentralisée l'exécution de la stratégie auprès d'un comité qui regroupe les lignes d'affaires.</p>		
---	--	--

Annexe 12 : Autres éléments déterminants pour une exploitation éthique des données tels que proposés par les gestionnaires

Autres termes proposés par manager		
L'analyse comportementale peut rapidement mener à des dérives éthiques. On peut imaginer par exemple cibler des comportements découlant d'une maladie. Chacune des utilisations devrait faire l'objet d'approbation d'un comité éthique.	Gouvernance des données	Acceptées
Les moyens utilisés pour extraire l'information.	Responsabilité/Imputabilité	Actuelles
Discrimination	Équité	Consentement
Absence de biais	Culture:	propriété
Portée	Stratégies:	Imputabilité
Autonomie	Normes:	Discrimination
L'intention		
Supervision par une partie indépendante	Règles gouvernementales / standards de l'industrie	Gouvernance des données

Annexe 13 : Certificat éthique pour le projet de recherche



Sherbrooke, le 23 novembre 2018

Mme Chimene Flore Fadje Ngayap
ÉCOLE DE GESTION (études)
Université de Sherbrooke

N/Réf. 2018-1797/Fadje Ngayap

Objet : Approbation finale de votre projet de recherche

Madame,

Le Comité d'éthique de la recherche – Lettres et sciences humaines a reçu les clarifications ou les modifications demandées concernant votre projet de recherche intitulé « **Études des éléments qui doivent être pris en compte de manière prioritaire dans une entreprise lors de l'établissement (ou de la mise en œuvre) d'une code de déontologie et d'éthique dédiée à l'intelligence d'affaires dans un contexte d'exploitation des données** ».

Les documents suivants ont été analysés :

- Formulaire de réponse aux conditions (F20-2519)
- Projet de recherche (projet-ethique-BI-Ngayap_V3.docx) [date : 13 novembre 2018, version : V3]
- Recrutement (projet-ethique-BI-lettre-invitation-Ngayap_V3.docx) [date : 13 novembre 2018, version : V3]
- Outil de collecte des données (projet-ethique-BI-questionnaire-Ngayap_V3.pdf) [date : 13 novembre 2018, version : V3]
- Formulaire d'information et de consentement (projet-ethique-BI-formulaire-consentement-Ngayap_V4.docx) [date : 22 novembre 2018, version : V4]

Le comité a le plaisir de vous informer que votre projet de recherche a été **approuvé**.

Cette approbation étant **valide jusqu'au 23 novembre 2019**, il est de votre responsabilité de remplir le formulaire de suivi (formulaire F5-LSH) que nous vous ferons parvenir annuellement. Il est également de votre responsabilité d'aviser le comité de toute modification au projet de recherche (formulaire F4-LSH) ou de la fin de votre projet (formulaire F6-LSH). Ces deux derniers formulaires sont disponibles dans Nagano.

Le comité vous remercie d'avoir soumis votre demande d'approbation à son attention et vous souhaite, Mme, le plus grand succès dans la réalisation de cette recherche.

M. Olivier Laverdière
Président du CÉR - Lettres et sciences humaines
Professeur au département de psychologie
Faculté des lettres et sciences humaines

c. c. Vice-décanat à la recherche

Directeur ou directrice de recherche (le cas échéant)

Service d'appui à la recherche, à l'innovation et à la création (le cas échéant)



Approbation du projet par le comité d'éthique de la recherche suite à l'approbation conditionnelle
CÉR Lettres et sciences humaines - Université de Sherbrooke / Tél. : 819-821-8000, poste 62644 / cer_lsh@usherbrooke.ca

1 / 1